

Glimmers on arms control

The ending of the arms control season is not as depressing as it seems. Geneva has been unproductive, Madrid a pleasant surprise.

ARMS control is indeed a fitful process: who would have thought that the best news, or at least the most tangible development, to emerge from the past six months of confusion should have come from the apparently endless negotiations within the framework of the Helsinki treaty of 1975 (called the Helsinki Final Act because it is not binding)? That, however, is what happened last week when, after more than two years of negotiations consisting more often than not of recriminations, the two power blocks with interests in West-Central Europe appear to have agreed at Madrid on a number of sensible compromises with each other.

The document that the states concerned will sign later in the summer contains some high-sounding principles about the freedom of individuals to make choices for themselves — declarations so general that only the most boorish would find it hard to subscribe to them — but which have been accompanied by a few exit visas for a number of Soviet citizens who have not hidden their unwillingness to stay (and who camped out in the United States Embassy in Moscow to make their point). Similarly, it seems to have been agreed that there are sensible things that can be done to reduce the risk of surprise attack, or conventional war without forethought, in Central Europe. There is to be a conference in Stockholm next year to carry further the provisions of the Helsinki Act (which requires each side to tell the other of manoeuvres within 150 km of the European border). This is not much to boast about but it is much better than nothing.

By comparison, the negotiations at Geneva from which much was expected at the beginning of this year have so far produced very little. The negotiation on missiles of intermediate range in Europe, which last week went into recess for the summer, appears to have degenerated into a bilateral restatement of conflicting positions. The Soviet Union's insistence that British and French nuclear weapons must be counted in any balance sheet is known to be unacceptable not merely to the governments concerned but also to the council of the North Atlantic Treaty Organization (to which France does not even belong). This does not imply that there is no framework within which these nuclear weapons will be counted; indeed, the Soviet Union must know that, with a little guile, it could persuade the United States to allow for them numerically without mentioning them explicitly, in the Salt II treaty (still unratified).

Inevitably, the Soviet Union can also complain that the United States position, based on the argument that the capacity of the Soviet SS20 force must be matched by the "modernization" of nuclear weapons in Western Europe, is similarly inflexible. The error in that argument is that the SS20s are a new element (whose number has been increasing even while the argument has raged) which present novel problems for the defence of Western Europe. It is not the Cuban missile crisis of twenty years ago all over again, but there are elements of similarity. When the negotiations began, it was widely thought in the United States that they would drag on until the first United States missiles were about to be deployed in Europe. As events have turned out, the Soviet Union probably has more to gain by waiting until the process has been under way for some months before softening its negotiating line, counting on the political discomfiture that will be caused by those opposed to nuclear weapons in any guise. The difficulty is that negotiating partners who have restated their position for the umpteenth time cannot then easily retreat into flexibility. The solution, for both

sides, is to recognize what has been apparent from the start — that the distinction between intermediate and strategic missiles is artificial, and should promptly be abandoned.

There, curiously enough, the past six months appear to have seen some progress. The two sides are at least talking about numbers, and have not abandoned the principle declared at the outset that the objective is to find some way of reducing the numbers of strategic weapons deployed on each side. (Each side's numbers is for the time being unacceptable to the other, but that gap could be bridged.) What seems, however, still to be lacking is a clear agreement on the functions of different kinds of weapons systems — aircraft (the most vulnerable), land-based missiles (accurate and less vulnerable) and submarine-based missiles (less accurate, even less vulnerable). Since the objective of the whole exercise is to find some way of regulating for stability, there is much to be said for relying on submarines rather than on land-based missiles (a proposition rejected by the Soviet Union some months ago, but which could be dusted off and tried again). But the best hope for these negotiations is that there should be a planned programme of reductions, with opportunities at pre-determined stages for reconsidering what happens next. Why not settle now for a one per cent reduction of existing forces and decide where to go next a year from now? Purists would say that such a process would never yield a final result, something to enable everybody to sleep easily at nights. But that is an illusory objective. Nuclear weapons will not go away. The best hope is that those who possess them will keep talking to each other.

This is not mere whistling to keep one's spirits up. That, as it happens, is how most existing arms control agreements have been reached. Apparently endless talk and the frustration that it engenders have eventually made agreement not merely possible but plainly prudent. This, it must be hoped, will be the outcome of the technical studies of a chemical weapons treaty and a test-ban treaty supervised by the Committee on Disarmament (also at Geneva), which plans to keep on working until mid-August. It would be too much to hope that both these projects will succeed in the near future. That either one or other might should be enough to keep the momentum going — and to offset last week's ambiguous decision by the US Senate that there should be a little more money for the production of chemical munitions. □

Paying for pollution

Environmental protection is not an absolute good but unavoidably an economic cost.

WHERE should the costs of environmental pollution lie? In the week in which Mr William Ruckelshaus, administrator of the US Environmental Protection Agency, has presented those living near a metal smelter in the state of Washington with the stark choice between their jobs and (possibly) their health (see page 200), in Britain the House of Lords Select Committee on the European Communities has published a sensible if inconclusive study of the "polluter pays" principle on which, in theory, most European pollution law is founded. Each case illustrates that there is a long way to go before pollution policy will be properly informed by rationality.

Ruckelshaus's conundrum has an old-fashioned quality,



belonging to the nineteenth century. What he appears to be saying is that the cost to the smelting company of reducing airborne concentrations of arsenic to the point at which people living nearby would not overtly be damaged might put the operating company in such financial difficulties that it could be compelled to shed jobs or even to close. So the local people, many of whom work at the plant, should be given a chance to say how they wish a balance to be struck. Ruckelshaus's main point, that too many people ignore the economic costs of pollution abatement, is fair enough, as is the common belief that absolute freedom from risk is possible. But he seems to have picked an unfortunate example on which to make his stand. For the people of Tacoma, Washington, are apparently being offered the chance to trade off between an economic benefit (job security) and a risk to their own health and that of their families. It is as if the workers in a nuclear plant were given an opportunity to cheapen the cost of their employer's operations by accepting high doses of radiation in return for extra payment, \$1,000 a rem, or something like that.

That there should be a market in people's willingness to shoulder environmental risks sounds logical enough, but this is the wrong market. There are two important objections to local dealing on pollution standards. First, it must be supposed that a community that now agrees to shoulder some likelihood of risk will retain the right to change its mind at some time in the future — hardly the best economic environment in which to operate. But the market in environmental hazards now proposed cuts across the more important pollution market — that arising from the principle that if the same rules apply to all polluters, and if the costs of complying with them are genuinely a burden on their trade, the long-term consequence will be that polluting enterprises will be located where the nuisance they cause can be most easily and most cheaply contained. This is the underlying national economic benefit of the principle of "polluter pays" — not to be mistaken for the kind of retributive licence it is supposed by extremists to be.

So how should the principle be applied? The ideal is that what the polluter pays (either by installing and operating equipment for the purpose or by payments to some public authority) should be related to the true marginal cost of pollution abatement. Properly calculated, these costs would increase faster than the gross amount of pollution, thus providing polluters with an economic incentive to be modest in their demands on the environment. The House of Lords study suggests that European practice, which is a long way ahead of the United States, falls a long way short of that. Too often, public authorities levy charges on polluters that provide no economic incentive (so much a unit of Biological Oxygen Demand for treating water, for example) and too often they have come to regard pollution charges as a source of revenue or as a means of subsidizing pollution abatement in other industries. (Even so, the committee does not share the opinion of British chemical manufacturers that European subsidies entail unfair distortion of what should be free trade.) So Europe believes that polluters should pay but so far has found only crude ways of making them do so.

Political chauvinism apart, three almost technical reasons account for the delay. First is the problem of sharing the absorptive capacity of some regional environment among the claimants on it. Should the whole cost be shouldered by the possibly efficient latecomers? This is what Gerard Hardin calls the problem of the commons. What can be done in the United States to reformulate pollution regulations in terms of the quality of the environment, which is what matters, rather than emission standards? And what can be done to aggregate the effects of several kinds of pollution so as to yield a unified index of, say, the hazards to human health that might then be accurately set off against a public understanding of what risks are acceptable? Ruckelshaus is entirely right to say that, ultimately, some such trade-off is necessary (and is already implicit in any attempt to regulate pollution). The truth is that freedom from pollution is a public good which must be a public purchase, as is elementary education for example. Local options may be easier to understand, but are diversionary. □

Trials on mistrial

A British anti-abortion pressure group is trying to ruin a well-planned study. It should stop.

WHILE some of the brickbats thrown at the British Medical Research Council over its controlled study of the efficacy of vitamin supplements in the prevention of spina bifida births (more generally, of congenital neural tube defects) have been merited, the council does not deserve the latest insult thrown at it — a publicity campaign by the Society for the Protection of Unborn Children whose declared objective is to bring the trials to a halt. The society is best known for its opposition to abortion, but last week it announced a plan to distribute half a million leaflets denouncing a programme which it calls "repugnant and immoral". Cleverly, the society intends to make its propaganda more effective by means of what it calls "pickets" stationed outside retail pharmacy stores owned by the company (Boots) supplying materials for the trial rather than outside the offices of the research council or the laboratories housing those planning the work that lies ahead. Credit for guile does not, however, excuse the exaggerations and misrepresentations of which the society's case consists.

The plan of the trial is now familiar. For many years, it has been suspected that congenital neural tube defects are a consequence of maternal vitamin deficiency. Neither biochemical investigations nor the less than fully controlled prospective studies so far carried out have, however, shown whether the deficiency is that of vitamins in general or of folic acid in particular. A series of studies coordinated by Professor Richard Smithells of the University of Leeds and involving women at several British centres has, however, shown that a vitamin preparation including folic acid appears to have a significant effect in reducing the chance of recurring neural tube defects in births to women to whom defective children have already been born. (Significance is judged by comparison with recurrence rates among women previously afflicted in this way but not taking part in the trials, usually because they were found already to be pregnant for a second time.) Three important issues are as yet unclear — the relative importance of folic acid and other vitamin supplements, the effectiveness of one, the other or both and the possible side-effects of folic acid given in comparatively large doses (4 mg a day) to large numbers of women. For it is accepted that a decision to use vitamin supplements of any kind will require that these should be offered to all potentially pregnant women, or all such women living in high risk areas. It is unthinkable that such procedures should be introduced, or that the use of a particular vitamin supplement should be publicly endorsed, without first being able to assess the benefits that might ensue, and the risks.

The Society for the Protection of Unborn Children has the benefit of an inner certainty on these difficult issues. It says that the trials are "totally unnecessary", protests at the use of a placebo (called "dummy pills") and asks why the Smithells regimen, "widely recognized throughout the world", should not promptly be introduced to Britain. The basis of this opinion is, however, clouded. The society's national director, Phyllis Bowman, seems not to have a formal mechanism for gathering medical advice but, on the telephone last week, had to rely on newspaper cuttings for her assertion that Professor Smithells, Professor J.H. Edwards (Oxford) and other medical people are "against the trials". That there are problems, and controversy about those problems, is not disputed. Some indeed consider that the case for the Smithells regimen is strong enough. Others are aware of the special character of this trial, which differs from most controlled clinical trials in that those involved are not manifestly ill (but their offspring may be). If the society had chosen to take an informed part in this debate, nobody would have complained. Its attempt to ruin the trials by picketing pharmacy stores, playing on the anxieties of pregnant women in the process, is discreditable. If the society wishes to keep its reputation untarnished it should acknowledge the shabbiness of what it is now about and call off its "campaign". □

US national laboratories

Packard report urges more autonomy without change

Washington

GOVERNMENT-owned laboratories in the United States should be freed from civil service pay codes so that they can recruit better scientists, and laboratory directors should be given much more control over their own budgets. These are the two main recommendations of a slender report, published last week after a year of work by a committee of the White House Science Council, which is otherwise virtually a carbon copy of earlier studies of the national laboratories published over the past year.

Chaired by David Packard of the Hewlett-Packard Company, the committee was created in March 1982 by presidential science adviser George Keyworth. It was asked at the inaugural meeting of the newly-established White House Science Council to take a "fresh look" at the 755 federal laboratories, which had not been the subject of a comprehensive study for more than two decades. The laboratories, owned by half a dozen separate government departments and occasionally managed by university or private contractors, account for about a third of the whole federal science budget. In 1984 they are expected to spend some \$15,000 million.

Although clearly one of the most important committees established by the science council, the Packard group decided against a comprehensive report and chose instead to set out a small number of general management principles for the complex of laboratories. The committee visited only 16 laboratories, eight of them belonging to the Department of Energy (DOE), and relied heavily on existing surveys of their work. Unsurprisingly, its conclusions are remarkably similar to those of a study of DOE laboratories published last September by the department's research advisory board (see *Nature* 12 May, p.101).

Like the research advisory board, the White House committee complains that many of the federal laboratories have diversified too much and need to be given clear and specific missions. It says that much of their work — on engine designs, batteries, fuel cells and renewable energy, for example — could have been done as well or better by the universities or industry. In the case of the DOE laboratories, a surge of work on alternative energy sources, followed by severe cut-backs, had left several major institutions without clear missions.

DOE is also singled out for being more prone than other departments to meddle excessively in its laboratories' detailed management. The Packard committee wants to give laboratory directors more independence by letting them spend

between 5 and 10 per cent of their budgets on projects of their own choosing. But they would also be made more accountable, by the establishment of external oversight committees and a greater use of peer review to monitor research quality.

Federal laboratories that are operated as well as owned by the government have a specially difficult time trying to recruit and retain scientific staff of a high calibre, the report maintains. Because they are tied to rigid civil service rules on pay and promotion, the laboratories cannot always reward scientific performance when it is not linked with managerial responsibilities. The committee wants scientific and engineering staff at these laboratories to be exempted from civil service salary policies.

The report claims to have identified

"several serious deficiencies" in the federal laboratory system, and says laboratories that are no longer fulfilling an important mission should be allowed to close. It makes no mention of specific laboratories, but at a press conference last week Mr Packard said the weapons laboratories at Lawrence Livermore, Sandia and Los Alamos should do less research on energy and more on weapons. And he said Lawrence Berkeley Laboratory, having lost much of its eminence in high-energy physics to Fermilab and the Stanford Linear Accelerator, might wish to define a new mission based on materials research.

At the same press conference, Dr Keyworth gave no hint of disappointment with the committee's work; it had been favourably received by President Reagan and had already been the subject of discussions between the White House science office, the Office of Management and Budget and the heads of government department which maintain federal laboratories. He said the administration regarded implementation of its recommendations as a "high priority". **Peter David**

Foreign relations

Silence to save shame

Washington

THE Reagan Administration was unable to refrain from patting itself on the back last week when it sent Congress its annual report on scientific relations with other countries. A covering letter from President Reagan pointed out that for the first time, scientific cooperation had been on the agendas of the international economic summits at Versailles and Williamsburg. The letter also noted that new scientific agreements have been concluded with India and Brazil and an old one, with China, considerably expanded. What the letter did not say was that the Office of Management and Budget (OMB) wants the report it accompanied — the only detailed study of how science affects US diplomacy — to be the last ever published.

Congress compelled the State Department to produce the annual report by writing a new section — Title V — into the Foreign Relations Authorization Act in 1978. The State Department was expected to provide a careful analysis of the part science and technology play in relations with other countries. Since the Title V reports began to emerge in 1980, however, they have generated more heat between Congress and the State Department than shed light on US thinking on international scientific relations.

Last year, the House of Representatives Foreign Affairs Committee was so dismayed by the vagueness of the report that it printed it with an unflattering critique provided by the Congressional Research Service. According to this agency, the report avoided candid discussion of the foreign policy or diplomatic content of most of the

activities described. Thus the withdrawal of the National Aeronautics and Space Administration from the International Solar Polar Mission was reported without mentioning the serious diplomatic repercussions, and a discussion of transboundary pollution between the United States and Canada did not refer to the impact of the acid rain debate on relations between the two countries.

This year's report does little better. A section describing the role of the United States at the United Nations' 1982 Unispace meeting in Vienna says the United States was able to demonstrate its preeminence in space applications, but does not hint that the meeting was widely viewed as a diplomatic rout for the United States (see *Nature* 21 April, p.646).

Congressional critics of the quality of Title V reports want to see them improved, not eliminated. It has therefore come as a surprise to the House Foreign Affairs Committee and the Committee on Science and Technology that OMB wants the Title V reporting requirement deleted altogether. The Title V reports survived the administration's purge of annual reports in last year's Congressional Reports Reduction Act, but OMB now claims that they merely duplicate the science and technology reports produced by the National Science Foundation.

Congress is likely to disagree. John Holmfeld, a staff official on the Science and Technology Committee, said most members regard Title V as an important way to remind the State Department of the scientific dimension in international relations. **Peter David**

Air quality standards

Arsenic and jobs trade-off

Washington

LAST month Mr William Ruckelshaus, the new administrator of the Environmental Protection Agency (EPA), called for a revolution in public attitudes towards the management of health risks. Addressing the National Academy of Sciences, he complained that the public and Congress expect an unrealistic degree of certainty about the health risks posed by toxic substances. Both would have to learn that life now took place in a "minefield of risks" and that the benefits of eliminating them must be weighed against the economic costs.

Last week, Ruckelshaus unveiled an almost perfect demonstration of how this new philosophy could be put into practice. Forced by a court order to publish a safety standard for the emission of inorganic arsenic, EPA is asking the residents of Tacoma, Washington, to play a major part in deciding between maximum safety and the closure of the local ASARCO copper smelter. The smelter emits about 282 tonnes of arsenic a year, but also employs more than 500 people in a city where unemployment is high.

EPA is compelled under the Clean Air Act of 1970 to set emission standards for hazardous air pollutants and, as the result of a court action brought by the state of New York, was given until last week to set standards for arsenic, which has been consistently linked in recent studies with skin and lung cancer in humans. EPA met its deadline, but made it clear to the residents of Tacoma and other areas affected by arsenic emissions that no standard can result in absolute safety.

In its draft standards, EPA claims that the Clean Air Act's requirement that standards must provide an "ample margin of safety" could not apply to carcinogenic chemicals which might produce effects at any level of exposure. The agency therefore proposes to make it plain to the residents of Tacoma that the proposed emissions standards could not eliminate all health risks posed by the smelter.

The EPA standard would require the ASARCO plant, the only smelter in the United States to use ore with as much as 4 per cent arsenic, to apply the best available technology to reduce the level of emissions. The measures would cost ASARCO about \$4 million, and would reduce annual emissions from 282 to 134 tonnes a year. "Fugitive" emissions, which pose most risk because they are released close to ground level and not through a high stack, would be reduced by 82 per cent.

According to the agency, the standard would reduce the cancer risk to the most exposed individual from about 9 per cent at present to 2 per cent and the risk for the thousand most exposed people from about 1 per cent to 0.2 per cent. Although only EPA is empowered to decide whether these

risks are acceptable, Ruckelshaus said last week that the residents' views would be a major factor: "We must ask them if they are willing to accept certain risks associated with exposure to low levels of arsenic".

If they are not, the residents of Tacoma may have no alternative but to force the smelter out of business. In its draft regulations, EPA maintains that further reductions in smokestack emissions would do little to reduce the risk. One alternative would be to compel ASARCO to use ore with a smaller arsenic content but, the

Radiation exposure

Japanese veterans cleared

Washington

US ARMY veterans stationed in Japan near Hiroshima and Nagasaki shortly after the cities were destroyed by atomic bombs in 1945 appear not to have suffered from an abnormally high incidence of the bone cancer multiple myeloma, the National Research Council (NRC) has concluded. A study* prompted by complaints from veterans' organizations identified only nine such veterans with the disease — at the bottom of a range of 9 to 29 cases expected statistically.

The findings come as little surprise to NRC, which maintained in a report two years ago that the low level of radiation to which US troops were exposed — about a tenth of a rad — suggested that the incidence of cancer could not be ascribed to exposure to ionizing radiation in the vicinity of the two cities. But the finding does contradict anecdotal evidence collected by the National Association of Atomic Veterans (NAAV), which says multiple myeloma occurs with unusual frequency among veterans who served in the area.

NRC conducted the survey to help the Defense Nuclear Agency to decide whether a full-scale epidemiologic study of multiple myeloma was warranted. A national advertising campaign and the establishment of a free telephone service produced 687 veterans who claimed to have been stationed near Hiroshima or Nagasaki. An initial list of 19 possible cases of multiple myeloma was whittled down to nine confirmed cases.

The report acknowledges that not all cases of multiple myeloma among the veterans are likely to have been identified and that it is therefore "conceivable" that the incidence of the disease is indeed abnormally high. But it says there is no evidence to support such a proposition. Studies of Japanese survivors exposed to less than one rad have found no difference in the frequency of multiple myeloma compared with the general population.

Since multiple myeloma has been the disease most cited by veterans who claim to

agency warns, it might prefer to close the plant altogether.

ASARCO has already announced that it is willing to spend the money necessary to comply with the proposed standard, but not whether it will close the smelter if EPA insists on the use of ore with a lower arsenic content. Meanwhile, EPA's attempt to make the public decide for itself on the level of acceptable risk has drawn mixed reactions. Ruth Weiner, professor of environmental studies at the nearby Western Washington University, said it was up to EPA to protect the public health, "not to ask the public what it is willing to sacrifice not to die from cancer".

Peter David

have contracted illnesses as a result of service near Hiroshima and Nagasaki, the NRC findings provide little support for a more detailed study of the troops involved. The council is, however, carrying out an extensive mortality study of veterans who participated in the atomic tests conducted by the United States in the Pacific and in the Nevada Desert.

Peter David

**Report of the Committee to Evaluate Cases of Multiple Myeloma Among Hiroshima/Nagasaki Veterans, National Research Council, Washington.*

No Dutch treat

Waalre, The Netherlands

UNIVERSITIES and teaching hospitals in the Netherlands will be on the average 8.3 per cent poorer over the next four years, with their budgets cut by Dfl.317 million (£70 million). Some 50 subject areas at the 13 Dutch universities will disappear and between 5,000 and 6,000 people, some of them part-time, will lose their jobs. Large sums will, however, be invested in new activities such as computer technology.

These changes are the consequences of a reduction in government spending announced by Mr Wim Deetman, minister of education and science, in October 1982. The universities were asked to make suggestions, which they did in May this year, and in the surprisingly short time since then the minister has decided on the more drastic cuts now announced.

Few faculties have been spared. Departments from different universities and different cities have been amalgamated. The three technical universities have come out best. A scientific advisory committee advised the minister to close a complete medical school in Amsterdam, which caused an uproar in medical circles and resulted in a reprieve.

It is too soon to know what the result will be for the quality of academic education and research, but there seems to have been little consultation or thought for the long-term effects.

Casper Schuurin

US National Science Foundation

Policy on minorities to change

Washington

SENIOR managers of the National Science Foundation (NSF) are battling against the calendar — and against some deeply entrenched personal convictions — to implement a dramatic change in the foundation's granting philosophy. A committee of assistant directors has been given until October to devise a new management plan to reconcile NSF's belief in selecting research proposals on scientific merit with its obligations to earmark some money for special groups such as women, minorities and international partners.

Until now, NSF has juggled the two roles by keeping them organizationally distinct. The heart of the foundation consists of six discipline-based research directorates, which support proposals chosen through peer review and judged on purely scientific grounds. A substantial number of special programmes, many of which are mandated by Congress and fit no particular research discipline, are grouped under a smaller Directorate for Scientific, Technological and International Affairs (STIA).

Dr Edward Knapp, NSF's director since the end of 1982, has opened a managerial Pandora's box by making the ending of this dual system one of his priorities. He maintains that the existing arrangements give NSF the worst of both worlds. Outsiders assume, often unfairly, that research sup-

ported through the specially directed programmes is scientifically inferior. Meanwhile, the existence of the directed programmes enables research directorates to take an excessively narrow view of their own interests.

Knapp told Congress earlier this year that in future the directed programmes would be shifted into the mainstream of the foundation's work. The programme officers of the research directorates would be expected to consider issues such as the support of undergraduates, minorities, women scientists and international partnerships when they were forming general plans for their disciplines.

The proposed reform has provoked a predictable barrage of protest, including some noisy criticism from Congress and more muffled arguments from within NSF itself. The House of Representatives Science and Technology Committee complained that the changes were being rushed through too rapidly and without enough planning and that the new arrangements would lead to the directed programmes simply being absorbed within the larger and less sympathetic research directorates.

Congress as a whole, however, did not support the committee's doubts. Under the 1984 budget, NSF will be allowed to begin to transfer some money out of STIA's budget and into the research directorates where, in theory, it can be spent on directed programmes. But the precise details of the new plan have not been finalized, and there are signs that disagreements within the foundation itself may result in less far-reaching changes than those originally envisaged.

One area in which Knapp's reforms are expected to be toned down is in STIA's international programme. In his initial explanation of the new approach, Knapp told Congress that STIA would relinquish its authority for peer-reviewing and selecting international research projects. STIA would remain responsible for servicing NSF's bilateral arrangements with foreign scientists, but half of its money — some \$5 million — would be transferred to the research directorates to distribute under arrangements similar to those used for domestic grants.

In a telephone interview, Mr Tom Ubois, the NSF assistant director devising the new management plan, said the foundation may decide against so radical a change. Many officials had questioned whether the research directorates could develop the expertise needed to cope with the special granting procedures that applied to some 30 bilateral international agreements operated by STIA. As a compromise, STIA might be allowed to retain its control over grants while allowing the research directorates more say in the planning of international agreements. **Peter David**

AIDS

Moral Majority intervenes

Washington

DR Jerry Falwell, the president of the Moral Majority whose Sunday morning sermons are televised and watched by millions of Americans, said last week that the acquired immune deficiency syndrome (AIDS) was a divine punishment visited on homosexuals for breaking the laws of nature and of God. And he accused the homosexual community in the United States of using its political influence to prevent the government from acting more quickly to stop the spread of the epidemic.

At a press conference on AIDS convened by the Moral Majority, Dr Falwell accused the federal government of failing to act against AIDS as rapidly as it had acted against legionnaires' disease, toxic shock syndrome and the poisoning of some capsules of extra-strength Tylenol last year. He



said federal health authorities should immediately close down all public bath houses, prevent homosexuals from donating blood and issue "firm guidelines" to nurses and other paramedicals responsible for handling AIDS victims.

Arguing that the sole purpose of public bath houses was to provide a setting for promiscuous homosexual sex, Dr Falwell said their closure would be analogous to the closure of some public swimming pools during the polio epidemics of the 1950s. And he claimed that while the spread of AIDS could also be checked by asking blood donors to answer extensive questionnaires detailing their sexual proclivities, Congress had shied away from such measures for fear of antagonizing the homosexual community.

Opposition to homosexuality and promiscuity, together with abstinence from alcohol, are major elements in the official programme of the Moral Majority, but Dr Falwell insisted that he felt great compassion for AIDS victims and supported more research to find the causes of the disease. Research, however, had to be accompanied by firm preventive measures to stop AIDS becoming an "uncontrollable plague" and spreading to normal Americans. **Peter David**

Axeman hesitates

BRITISH universities seem likely to suffer as a consequence of a £30 million cut in expenditure for the current financial year at the Department of Education and Science (DES). Details of where exactly the axe will fall have yet to be decided.

There will be no change in university student numbers for the coming academic year, but, in addition to the universities themselves, research councils and institutions supported directly by the department — such as the Open University — seem vulnerable. However, Sir Keith Joseph, Secretary of State for Education and Science, is said to be anxious to protect "important scientific research" and, in particular, his recent "new blood" initiative — which sought to compensate for some of the ill-effects of earlier cuts. The latest cut amounts to 1 per cent of DES's science and education expenditure.

Mr John Acker of the Association of University Teachers says that any further reductions in universities' budgets would be disastrous for research and for staff morale. He is pinning his hopes on a pledge apparently made by Mrs Thatcher during the general election campaign that spending on higher education in and after 1984–85 would be held at 1983–84 levels. **Tim Beardsley**

UK biotechnology

A public for Celltech?

To cap a flurry of recent announcements from Britain's leading biotechnology company, Celltech, Gerard Fairtlough, its chief executive, said last week that it is "entirely possible" that the company will go public within five years. In the meantime, it needs to find new admirers among financial institutes both to provide additional capital and to replace the British Technology Group (which next week will announce the formation of a new biotechnology company) in its measured withdrawal as the major shareholder.

Fairtlough sees no problem in finding new investors, and as evidence of Celltech's viability, he points to the advanced plans for setting up Boots-Celltech Diagnostics Limited and the opening this week of a Culture Products Division of Celltech.

The new division has the task of bulk production of monoclonal antibodies. Some, such as interferon and ABO blood group monoclonals, are for Celltech to sell. Others are on contracts, about six of which have been negotiated with companies that include ICI and Centocor, at around £500 per gramme of antibody for quantities of less than 1 kg.

The division has at present two 100-litre fermenters in which a batch of antibody-producing hybridoma cells can be grown every week or so. With a £100,000 grant from the British Government's Department of Industry, Celltech expects soon to be able to develop a pilot-scale continuous-culture fermenter in which hybridoma cells will survive and produce antibody for several months at a time.

It also claims that, contrary to many opinions, rodent hybridoma cells are perfectly capable of growing in suspension rather than in the protected environment of capsules or fibres and that the yield of antibody from such cells is several hundred microgrammes per millilitre, an order of magnitude greater than had been thought possible. Celltech is confident that human hybridoma cells will behave just as well. And it claims to be in the final stages of development of a serum-free medium in which to suspend the cells during fermentation.

Celltech's plan to make money out of the genetic mapping of microorganisms on contract is, however, going less well. The procedure was devised in Dr Sydney Brenner's Medical Research Council (MRC) laboratory in Cambridge, but its development into a commercial process is "much slower than hoped".

The future of the agreement whereby Celltech has first rights of refusal on any MRC invention in the area of recombinant DNA or monoclonal antibody research is also up for grabs. Fairtlough, unsurprisingly, would like it to continue and believes MRC has been delighted so far.

The agreement runs until 1985 but is now being reviewed. One crucial factor is whether MRC will wish to carry on. Another is whether Celltech will be able to secure first rights even in a circumscribed area when the National Research Development Corporation, originally a reluctant partner in the arrangement, and which collects a rent from Celltech to cover loss of benefit, is soon to lose its first claim on research council research. Most probably, the monogamous agreement between Celltech and MRC will turn into an agreement that is polygamous on both sides.

Peter Newmark

Satellite storm ahead

Washington

BRITAIN is adding its voice to the chorus of domestic criticism of the Reagan Administration's plan to sell the government's remote sensing and weather satellites to private business. A frosty memorandum spelling out the British attitude warns that the sale could damage scientific cooperation between the two countries and could have "serious repercussions" on international weather forecasting.

The memorandum, which has been made available to the congressional committees concerned, describes the free exchange of meteorological information between national weather services as a "fundamental principle" which should be safeguarded in any arrangement the United States makes to sell its satellites to private operators. If the United States began to charge for its satellite observations, the British note claims, other countries would follow suit.

Britain is also worried about the impact of a sale on an agreement to provide instruments for the United States' polar-orbiting meteorological satellites. The United Kingdom Meteorological Office already provides stratospheric sounding units for the American TIROS-N satellites in exchange for free data from the satellites. In recent months, the two countries have agreed in principle to extend the arrangement, with Britain providing an advanced microwave sounding unit.

The British memorandum expresses less concern about the fate of the remote sensing satellites, but says it is "doubtful" whether the time is ripe for the wholesale commercialization of such systems. Pointing out that individual British research laboratories are working as principal investigators on remote sensing with the National Aeronautics and Space Administration, it says the United Kingdom would be "very concerned" if the proposed changes led to a reduction in similar opportunities for collaboration.

Peter David

Spanish research

Discreet threat about jobs

Barcelona

THE newly appointed president of CSIC (Consejo Superior de Investigaciones Científicas, the Spanish science research council) has made an increase in CSIC personnel his "absolutely priority" task and intends to make it a condition of his continuing in his job. Other points that he stressed in his first speech were a contraction in the number of lines of research carried out by CSIC institutes and the need to consider the possible transfer of the CSIC institutes to the autonomous governments now established in the different parts of Spain.

Dr José Elguero Bertolini is the first CSIC scientist with considerable scientific experience abroad to have been appointed president of CSIC, a job traditionally given to a university professor. An organic chemist, he has spent most of his career in France. His scientific and international experience will be very useful in the present move to redefine the function of CSIC.

In the first declaration on science policy by a member of the government since it took office eight months ago, Mr José Maravall, the Minister of Education and Science, echoed in a speech to the Senate much of what Dr Elguero said. He confirmed the intention of the Socialist Government to increase public spending on research. He stressed the need to reorganize the structure of organizations concerned with science policy in Spain, to propose a number of priority research fields and to define the place for research in the new structure of the Spanish state, with its 17 autonomous governments. Four fields of research to be given priority in conjunction with Spanish industry are energy use of biomass, aquaculture, microelectronics and railway technology. Two broad fields have been chosen for priority financing in basic research: high-energy physics, which will be immediately affected by the recent reentering of Spain into the European Organization for Nuclear Research (CERN), and biotechnology.

The promised increase in public spending on research has already begun, and CSIC this year has 21 per cent more money. The government's general policy, however, is directed to reducing public expenditure, particularly on administrative staff, and if that policy is applied to research, the effect of the new funds would be very much reduced. The scientific staff of CSIC has not been increased for many years and it is now reaching an average age of 48, while many young scientists in Spain and abroad are anxious to join the new Spanish effort on science. CSIC is keeping its fingers crossed that its new president will have no reason to resign soon.

Pedro Puigdoménech

French research

Rigidity of science contracts

FRENCH researchers and technicians working for research councils such as the Centre National de la Recherche Scientifique (CNRS) may have to wait until September before they win the coveted security of "*Fonctionnaire*" (civil servant).

A new contract of employment granting this status, at present enjoyed only by university employees, was expected this month but has run into serious difficulties with the trades unions (which represent predominantly technicians).

Although the contract has been approved by the Prime Minister, Pierre Mauroy, and three other ministers concerned (industry and research, civil service and finance), the government may now give the unions time to produce a reasoned response. This will push the date of the final decree into the autumn unless the French administration follows in the steps of its predecessor, which took to publishing controversial legislation during the summer.

The unions are not alone in objecting to the new contract. Researchers who helped to draw up an earlier version in consultation with the previous research and industry minister, Jean-Pierre Chevènement, argue strongly that the present draft fails to do justice to the special nature of research careers. And directors of CNRS itself are also concerned. So if the present version is pushed through this summer, or even succeeds unchanged in the autumn, there is likely to be an unholy row.

The objections centre on the "exceptions" that have to be made to the usual civil service contract to accommodate promotion on merit and appointments to high-level research or technician posts from outside. When a working party at the National Colloquium on Research and Technology in January 1982 decided to seek *fonctionnaire* status for technicians and researchers, it assumed that such possibilities could be accommodated.

The draft contract thus provides that an entrant will be guaranteed a small increase of salary throughout his or her career, but that promotion to fundamentally new levels will be determined purely on the basis of performance. The underlying progression would avoid the problem that a technician, for example, can be "blocked" for ten years at one salary. One consequence is that people are then promoted more from compassion than on scientific judgement.

In 1982, the minister of the civil service was encouraging: he would welcome such a radical example to improve career development in other parts of the civil service, where appointments and promotion are determined by paper qualifications and rigid quotas.

However, the minister has since been advised that some 1-2 million other civil servants might demand exactly the same

rights and salary rises as the scientists and technicians, so that the revolution might go too far and cost too much.

Chevènement, apparently, managed to hold his ground for a time, but was already slipping when he resigned. The new minister of industry and research, Laurent Fabius, seems to have given up the fight. The result is that French researchers and technicians are being offered a more rigid draft contract than that already in existence.

Pierre Papon, director-general of CNRS, who saw the draft contract last

week, said on Monday that there was still "a degree of freedom" in the wording of the contract, and that "a kind of compromise" between the rigid civil service view and the earlier colloquium view was possible by September. The main problem is to be able to hire experienced people at the right pay and level, he said.

The chances of a slow underlying slope in salary, plus promotion on merit, was, however, "not completely clear". It is certainly not in the present text, he says. This problem is not so much for the scientists as for the technicians and engineers, says Papon — which is exactly where the unions come in and why French research could find itself in difficulties with technicians' strikes by the autumn. **Robert Walgate**

UK geothermal energy

Third hole lucky for Camborne?

THE British "hot dry rocks" geothermal project run by the Camborne School of Mines in Cornwall has been promised another £11 million by the Department of Energy. This will cover the last three years of a seven-year research programme. In spite of the difficulties encountered earlier this year in establishing strong fluid flow between its two existing 2,000-metre-deep holes, the Camborne team has thus succeeded with its argument that its first holes were essentially "wildcats" — against some scepticism within the Department of Energy.

The two existing holes, and attempts to connect them by "hydrofracking" and other methods, have taught the Camborne team "a lot of about the stress regime" deep in the Cornwall granite, according to the deputy project director, Dr John Beswick, last week. While the behaviour of the rocks is consistent with what is known of tectonic forces, the fracturing is more difficult to understand. Television observation reveals cracks at the holes themselves, but how the cracks spread between the holes is uncertain. Microseismic observations have suggested that hydrofracking had opened fractures beneath the holes, but — said Dr Beswick — the microseisms indicated movement below the holes, but not necessarily the existence of flow paths.

Some £3 million of the new £11 million will be spent on drilling a third 2,000-metre hole, always envisaged, and it now seems sensible that it should go a little below the bottom of the present holes into the zone of the microseismic events. A key problem now will be to determine exactly what that fracture regime is. That there are opened fractures is clear: "we can pump 180 litres per second of thick gels down a hole" said Beswick. The trouble is that, at present, relatively little of that volume comes back up the second hole. Camborne must arrange that it comes up the third hole to be drilled (or *vice versa*).

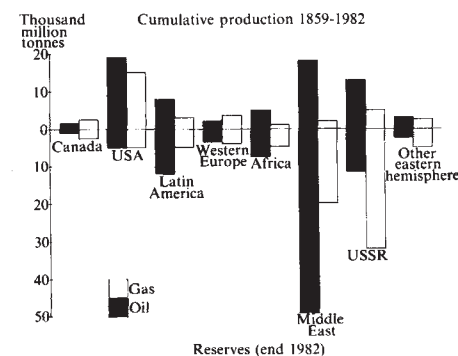
The cash will also be used to automate

certain tedious parts of the experiments and to prepare for a 6,000-metre well beyond the present seven-year programme. Some support for that may come from the European Commission, but Brussels is interested in supporting one of two new French projects in the interests of diversity.

The Camborne project is now seen as complementary to that at Los Alamos, New Mexico, which works at much higher temperatures (and lower depths) on the side of a volcanic caldera and which has encountered problems of shielding electronic equipment and short tool lifetime. The Camborne team argues that low-temperature sources (which are geographically common) may be more economical sources of geothermal energy.

Robert Walgate

World oil resources



In spite of the much increased production of crude oil from the Middle East, the United States remains the territory from which most oil has been produced in the past century. The chart (taken from *BP Statistical Review of World Energy*) shows cumulative production and proved reserves of petroleum and natural gas (at 1982) in different regions of the world. Criteria for the determination of reserves as "proved" vary from place to place.

Psychiatry and ethics

What cuckoo's nest next?

Vienna

THE Soviet resignation from the World Psychiatric Association (WPA), followed by the resignations of the Bulgarian, Czechoslovak and Cuban psychiatric associations, cast a shadow over the Seventh World Congress on Psychiatry here last week. These resignations did not, however, herald a general walk-out of the Socialist bloc as had been feared. Romanian and Polish psychiatrists attended the Congress but not the meeting of the WPA executive. However, the Hungarians came in force with more than forty people, mainly interested in the anomalously high suicide rate of Szeged county, while the East Germans made a spirited stand at the general assembly to keep "politics" out of the association.

The reason for the Soviet resignation from WPA was the well-documented allegations of political abuse of psychiatry to repress dissent in the Soviet Union. Unfortunately, at one press conference, the outgoing WPA secretary, Dr Peter Berner, gave the impression that the problem was simply a divergence of theory. In fact, the theory of "creeping schizophrenia" (a disease unknown outside the Soviet Union) promulgated by Dr Andrei Szeshnevskii is frequently cited by the Soviets to justify

"political psychiatry"; but, as Dr Anatolii Koryagin, now serving a seven-year prison-camp sentence for an article in *The Lancet* revealing the dark facts of Soviet political psychiatry showed, such theories are frequently a facile "justification" for morally untenable practice.

The Soviet resignation was accepted with "regret", and a motion from the United Kingdom Royal College of Psychiatrists was accepted with an overwhelming majority, assuring the Soviets that they were welcome to rejoin as soon as they can prove they have abandoned doubtful practices.

The issue of psychiatric ethics cut deeper than this political issue — or indeed, the converse allegations of malpractice and neglect in "black" mental hospitals in the Republic of South Africa. Behind the frankly commercial aspects of much of the conference the ethical issues of psychiatry constantly manifested themselves. In particular, papers were presented on the apparent conflict between patient confidentiality and the need for comprehensive epidemiological surveys, the need to relate psychiatric and psychotherapeutic treatment to the cultural background of the patient and the problems of patient consent in carrying out double-blind tests on new psycho-pharmacological agents. **Vera Rich**

UK ministry decrees drug price cut

THE Association of the British Pharmaceutical Industry (ABPI) said it was "shocked" after Mr Norman Fowler, Secretary of State for Social Services, last week persuaded industry representatives to accept a £25 million reduction in the National Health Service (NHS) drugs bill for the current financial year. ABPI

2.5 per cent less and freezing all prices until next April. ABPI members spent £400 million on research in Britain last year, and 60 per cent of their sales revenues comes from the home market; profits on sales to NHS amounted to £200 million.

ABPI says that the new cuts fall outside the scope of the Pharmaceutical Price Regulation Scheme, which sets individual profit targets — calculated on capital employed — for suppliers to NHS. But the association may have worse surprises in store. The House of Commons Public Accounts Committee recently suggested a much larger cut in the average profit allowed, and the Department of Health and Social Security is now conducting a "major review" of the whole subject. Last year, nine companies exceeded their profit targets, although less than £1 million of the total £33 million excess is likely to be reclaimed; the rest falls within a 10 per cent "grey area" over the targets that is meant to provide an incentive to efficiency and exports. This method of regulating profits has been criticized for penalizing those companies whose research efforts are more successful. The industry is also worried that Mr Fowler will make changes designed to encourage "generic" prescribing, recommended earlier this year by the Greenfield report but on which the government has not yet acted. **Tim Beardsley**



warned that "such repressive measures if continued or extended will damage investment confidence leading to a reduction of research activity".

Mr Fowler's request was part of a £1,000 million package of expenditure cuts and asset sales agreed by the Cabinet. The savings on the NHS drugs bill, now running at about £2,000 million per year, will be achieved by paying suppliers an average of

African famine

Can science help in time?

THE green revolution is still racing against the threat of famine, according to a group of leading agriculturalists meeting in London last week. Africa, for example, faces "mega-disaster" in 10–15 years and not in the "comfortable middle distance". The Interunion Commission on the Application of Science to Agriculture, Forestry and Aquaculture (CASAFA), an offspring of the International Council of Scientific Unions, is seeking to drum up serious interest in realistic practical questions which might nevertheless be of some assistance.

CASAFA, meeting to define its first research programme, showed its pragmatic approach by rejecting 57 out of a list of 64 project proposals. It is not that there is a lack of money. Sir Charles Pereira, a CASAFA member, said that the aid agencies are swamped with good intentions but not with good ideas and that one of CASAFA's objectives is to sort out "the bids based on pure optimism from those based in some fact".

The priority list agreed last week includes "allelochemicals" — low molecular weight compounds, on one case simply oxalic acid, which seem to be given off by some plants (including some rice species) to repel insects; the physiology of salt tolerance in rice; soil physics and aggregation of soils in rice fields; the link between tolerance to heat and winter hardiness in wheat; flower and fruit fall in potato; the symbiosis of *Frankia* species (nodule-forming nitrogenous bacteria) with *Casuarina*, a widespread tropical and semi-tropical tree whose many varieties can grow in conditions from desert to salt water and which burns fiercely; and the control of *Striga*, a parasitic weed now badly affecting Africa.

CASAFA's list is strictly a list of priorities. Other topics are also worthy of support. The physiology and pathology of aquaculture species is little known but important, especially because the protein content of fish does not denature as easily as meat. For Pereira, the large reserves of volcanic rock phosphate are a key issue because there are no facilities to test them in the tropics.

As for the continued dependence of the green revolution on fertilizers, the CASAFA group is pessimistic. High-cropping inevitably requires high feeding, and the genetic engineering solution seems a long way off, and anyway might absorb so much energy from the plant that crops would be reduced. So CASAFA pointed the finger only in one direction: at the energy of the flare-off from Arabian oil wells, which could provide 50–100 per cent of nitrogen-fixation needs of the developing countries if only they were fully used. **Robert Walgate**

Goitre in India

Iodine prophylaxis falters

New Delhi

A SURVEY by the All India Institute of Medical Sciences (AIIMS) in New Delhi has made the alarming discovery that one out of 25 babies born in the country's endemic "goitre belt" is hypothyroid and is thus destined to be mentally retarded. Some 40 million Indians have goitre and about 300 million live in the goitre belt which runs through 15 states south of the Himalayas. Considering that about one million births take place annually in this region, the AIIMS data suggest that India may be adding about 40,000 idiots every year to its growing population simply because they do not get enough iodine in their food.

The survey was sponsored by the Department of Science and Technology and carried out by a team led by Dr N. Kochupillai of the department of medicine at AIIMS. The survey covered three of the 16 goitrous districts in the state of Uttar Pradesh, where the incidence of goitre varies from 30 to 70 per cent. Nearly 5,000 children born every year in each of the goitrous districts are destined to be mentally retarded, Dr Kochupillai said. "What we found is alarming", he added, "goitre cannot be simply dismissed as a cosmetic problem."

The survey of neonatal hypothyroidism was made possible for the first time in India by a sensitive and specific radioimmunoassay developed at AIIMS for measuring



Incidence of goitre in India. Areas of endemic goitre are shown shaded.

picogrammes of thyroid hormones in cord blood samples collected on filter paper. Samples were collected by health workers during each delivery at the village primary health centres and sent by post to AIIMS for analysis. Three different thyroid hormones were measured and a diagnosis of hypothyroidism was made if all three indicated thyroid failure. The AIIMS group has so far analysed about 1,400 blood samples, of which four per cent are positive for hypothyroidism. In contrast, only one case of hypothyroidism was found among the 1,400 samples taken from Delhi and Kerala state, where goitre is not endemic.

The field survey by the AIIMS team also revealed a high incidence of cretinism, deaf-mutism, squints, stunted growth, hearing and speech disorders among goitrous patients. In some villages of Gonda district, the entire population is goitrous. Women hide their neck-swelling under multi-stranded necklaces, and girls have their throats painted with iodine concoctions sold by quacks. The IQ of children affected by goitre is so low in Katrasana-bajpur village that they hardly finish the fifth standard and 70 per cent drop out even earlier.

The prevalence of goitre in the Himalayan belt was well known in the early decades of this century. A controlled trial of the prophylactic value of iodized salt was mounted in the Punjab in 1954 and



Goitrous neck-swelling

both iodate and iodide were known by 1962 to be effective (see Ramalingaswami, V. *Bull. WHO* 49, 307; 1973).

The findings have come as a shock to the government, whose 20-year-old programme to provide iodized salt to the goitre belt has failed miserably due to logistics and the apathy of state governments. The village primary health centres do not list goitre as a health problem and people have not heard about iodized salt, says Dr C. Gopalan, former director general of the Indian Council of Medical Research and now president of the Nutrition Foundation of India. Uttar Pradesh state has yet to identify the sites for two iodization plants donated by the United Nations International Children's Emergency Fund (UNICEF) some four years ago. The control of goitre affecting 40 million people continues to be nobody's business.

The AIIMS team is now injecting iodized oil donated by UNICEF in selected populations in the study areas. The injections will dissolve the neck swelling but can do nothing to correct the irreversible mental damage that results when hypothyroidism is untreated in the first year of life. The team has also formulated a \$4 million scheme to administer iodized oil to six million goitrous women of reproductive age.

K.S. Jayaraman

US high schools

Models for innovation

St Louis, Missouri

TWO recent innovations, in North Carolina and Missouri, have relieved the general gloom about the quality of US public education. Both state-backed programmes aim to upgrade the quality of science and mathematics education by bringing high-school students more directly into contact with professional people.

North Carolina's model is its School of Science and Mathematics (NCSSM) at Durham. It is a residential public school open to 400 juniors and seniors from the state who have been judged to show exceptional promise. The state meets the cost of room and board as well as tuition. Classrooms and dormitories are housed in an old hospital complex. The school was begun at the urging of Governor James B. Hunt and has now graduated its second class.

Besides offering a challenging academic programme, the school puts its students into the laboratories of working scientists. It is within driving distance of four universities as well as Research Triangle Park. Most students choose to enter the school's mentor programme, in which they work on a project supervised by a scientist from one of the universities or the research park. Each student spends half a day each week in the laboratory and some work summers as well. Some professors attend the campus to help the school's highly trained staff.

The excellence of the teaching at the school is supposed not to be confined to the student elite chosen to attend. To upgrade science and mathematics education in other public schools in the state, NCSSM conducts workshops for teachers and brings in visiting teachers as fellows. At least one state, Louisiana, is ready to follow North Carolina's example and a further 20-25 are interested.

In a less ambitious programme, but one also likely to be copied, Washington University at St Louis is "recycling" scientists and engineers from private industry into science and mathematics education. The idea came from Carol H. Pauk, assistant director at the Graduate Institute of Education, and was stimulated by a programme of early retirement among the nearby Monsanto Company's professional staff. She persuaded the company to pay most tuition costs and to adjust the working hours of employees wanting to take the certification courses necessary to teach science and mathematics.

Pauk's initial concern that professionals switching careers in their 50s might not be able to adjust to the stresses of the classroom has not been borne out — perhaps because all those taking the course so far are married to teachers.

Karen Freeman

A nightmare of AIDS

SIR — In your timely article on acquired immune deficiency syndrome (AIDS) (*Nature* 2 June, p.377) you pertinently pointed out its possible relation to human T-cell leukaemia-lymphoma virus (HTLV). This is an apparently infectious human oncogenic retrovirus, recently found independently in Caribbean areas as well as in the United States and Israel on the one hand and among the south-western Japanese on the other hand. The various isolates proved to be similar in their molecular properties (*Nature* 302, 567; 1983). The Japanese virus strain, called the adult T-cell leukaemia virus (ATLV), can be activated *in vitro* from virus-free, peripheral-blood mononuclear cells of adult patients suffering from leukaemia¹, and the specific virus antibody was found in 25 per cent of the healthy population of the area where ATLTV is indigenous. Moreover, the virus could be produced by activation *in vitro* from the blood cells of apparently healthy people^{1,2}.

HTLV is similar to mouse Moloney leukaemia virus³ in its chronically oncogenic nature, and may be classified as a weak oncogenic retrovirus which does not contain an oncogene but may be integrated in the vicinity of a cellular oncogene of the host chromosome to cause leukaemia at a later stage of life. It was also shown in this connection³ that the inactive virus genome integrated into the chromosome in an artificially-produced chronically leukaemic substrain of mouse could be activated, through demethylation, when the Moloney leukaemia provirus DNA from liver of the non-leukaemic mouse of this substrain was cloned in bacteria³. In the past we have raised the question as to whether this last piece of evidence is relevant to the problem of safety in recombinant DNA work^{4,5}. While preparing the article⁴, I had a chance to point out to the Recombinant DNA Monitoring Committee in Australia that, contrary to its conviction, this kind of leukaemia retrovirus should be relevant to our safety concern because it might be related to the ATLTV then already known to science for nearly a year. The Australian Committee enquired about the problem of some overseas authorities on recombinant DNA, including GMAG in the United Kingdom. Apparently the overseas bodies reassured it that the problem might be dismissed without the need for any experimentation. In fact, you reported (*Nature* 302, 465; 1983) that most GMAG members might not be up to date in their knowledge of the subject on the one hand, but also that GMAG should be phased out on the other hand, because the concern about safety of recombinant DNA is also becoming out of date.

I now feel that the concern and uncertainty about AIDS at one end and the smug complacency about the safety of the

recombinant DNA work involving oncogenes and retroviruses at the other end are not logically consistent attitudes to be taken by the same scientific community, especially when there is a suggestion that the two ends may be linked by a common agent, the leukaemia retrovirus. May I quote from the two articles in *Nature* which I referred to above? (Italics are mine.)

"The popular *nightmare*, that genetic manipulation might, even by accident, create especially virulent microorganisms, was never more than a scare" (7 April p.465; and "... the unexpected appearance in the past few years. of this bizarre disease [AIDS] has provoked a sense of *nightmare*" (2 June, p.377).

A. SIBATANI

*CSIRO Molecular and Cellular Biology Unit,
North Ryde (Sydney),
NSW 2113, Australia*

1. Hinuma, Y. *et al.* *Gann* 73, 341-344 (1982).
2. Gotoh, Y., Sugamura K. & Hinuma, Y. *Proc. natn. Acad. Sci. U.S.A.* 79, 4780-4782 (1982).
3. Harbers, K., Schnieke, A., Stuhlmann, H., Jähner, D. & Jaenisch, R. *Proc. natn. Acad. Sci. U.S.A.* 78, 7609-7613 (1981).
4. Bartels, D., Naora, H. & Sibatani, A. *Trends biochem. Sci.* 8, 78-80 (1983).
5. Bartels, D. *Search (Sydney)* 14, 89-92 (1983).
6. Ada, G.L. *Search (Sydney)* 14, 92 (1983).

To be or not to be

SIR — In his interesting commentary on consciousness (*Nature* 3 May, p. 11), Professor Crook touches on the problem of defining living beings as distinct from man-made things. He correctly states that "all organisms share a self-referential cybernetic physiological system evolved to ensure self-maintenance at least until reproduction". This is a rather abstruse way of saying that "all organisms not only exist but also want to exist, that is, they have a self-directed will".

The operative cause for this is obvious: the subjective correlate of an assault on existence is pain, while the subjective correlate of the removal of such an assault is pleasure.

The ultimate cause belongs to the domain of metaphysics. As I stated before (*Nature* 228, 888; 1970) the problem appears insoluble, because we have known since Kant that the categories of being cannot be used to define Being as such, or, to put it another way, the will as "*Ding an sich*" cannot will anything but itself. Its choice is limited to the alternative: to be or not to be.

Here religion can transcend metaphysics by stating that: "All matter longs for life, all life longs for (higher) consciousness, all consciousness longs for God".

S.V. VAECK

*Laboratoire Central,
Administration de l'Industrie,
Ministère des Affaires Economiques,
Brussels, Belgium*

Decision-making

SIR — How to achieve "truly democratic" methods of choice between alternatives, a question raised by Professor R.M. May¹, transcends purely local problems.

Other examples of dissatisfaction with the performance of those whose business it is to take decisions, from the single voter to ministers or heads of state, are easy to find. There are two on the same page of a later issue of your journal², an editorial in a parallel publication³, another in a more specialized contemporary⁴ and so on.

Amongst many known reasons for failures to decide well, or to decide at all, are: failure to define the problem, failure to decide how to choose those who are to decide, confusion between the roles of the expert and the legislator, basic limitations in human mentation such as the lack of transitivity in statements of preference. Problems will continue to arise so long as decisions are reported solely in terms of their outcome, and the processes leading to them remain unexamined. Examples from many different fields show high correlations between final judgements. The majority of studies stop there, and so fail to show that very different strategies may have been used by individuals who "agree". They may use different sources of information, weight the same sources differently or combine them in different ways. The danger lies in their assumption that their strategies were identical because their overall judgements were in agreement. When they fail to agree on subsequent occasions, there will certainly be surprise and there may be disappointment, incomprehension or outright conflict.

It is not enough to show high overall agreement between managers and unions⁵ or between medical diagnosticians⁶ on a particular set of instances. To predict the extent of agreement in the future, and (if agreement is considered to be desirable) to increase it, it is essential to describe explicitly the judgemental strategies of those whose decisions carry weight in any size of community. Fortunately, a large body of research addresses such problems⁷⁻⁹, all tending to show that cognitive analysis is more likely to lead to happier committees and better decisions than game theory, second-guessing confrontation or agenda-less summit meetings.

C.R.B. JOYCE

*Beim Lindenbaum 15,
4123 Allschwil, Switzerland*

1. May, R.M. *Nature* 303, 16-17 (1983).
2. *Nature* 303, 270 (1983).
3. *Science* 220, 13 May (1983).
4. *Br. med. J.* 286, 1079-1081 (1983).
5. Balke, W.M., Hammond, K.R. & Meyer, G.D. *Admin. Sci. Q.* 18, 311-327 (1973).
6. Fisch, H.-U., Hammond, K.R., Joyce, C.R.B. & O'Reilly, M. *Br. J. Psychiat.* 138, 100-109 (1981).
7. Hogarth, R.M. *The Psychology of Decision* (Wiley, Chichester, 1980).
8. Hammond, K.R., McClelland, G.H. & Mumpower, J. *Human Judgment and Decision Making* (Praeger, New York, 1980).
9. Hammond, K.R. & Joyce, C.R.B. in *Drug Monitoring* (eds Gross & Inman) (Academic, London, 1977).

Intelligence, guesswork, language

H.B. Barlow*

A satisfactory definition of intelligence has never been found, and as a result it means different things to different people. What it is may remain too complex for a succinct definition, but the theory and practice of information handling have clarified what it does for us: it enables us to guess better, and the discovery of unexpected orderliness is the chief means of doing this.

SIXTY years ago E.G. Boring¹ pointed out the sorry fact that, for those who believed in tests, intelligence was simply what the tests tested for. This circular definition could have been escaped by using an objective, external, criterion for validating them, but in spite of much discussion about the nature of intelligence by the pioneers of mental testing no agreement was ever reached, and no such criteria were available. For a long time the testers do not seem to have worried about Boring's stultifying dictum, but a reassessment of the problem is now timely for two reasons. First, a glance at current literature on the sociology, psychology, and biology of intelligence reveals an unhappy confusion that must partly originate from disagreement about its nature. Second, theoretical understanding of the information processing operations that must underlie intelligence has increased to the point where it influences how we question the nature of intelligence: instead of asking "what is it?" we ask "what does it do for us?" Computers are our servants, but they now claim to be intelligent; even if they are not successful their claims may help us to clarify our own minds on the subject.

Disturbed intelligence

Attitudes towards intelligence vary a great deal among the testers themselves: some seem to be completely satisfied with current tests and apparently regard them as valuable and precise scientific measuring instruments, whereas others believe that the nature of intelligence is such that it is better to avoid altogether speaking of its measurement²; on this view quantitative tests are simply an aid to forming a judgement and a phrase such as 'subnormal intelligence' means just about what the man-in-the-street takes it to mean and no more. A revealing exchange about the role of psychological evidence in a court case^{3,4} brings out this clash between the desire to exalt the status of intelligence tests and the view that test results should be considered on the same level as other evidence. The court preferred the latter view, as will many others.

A third approach among the testers is

that of the cognitive theorists. Hunt⁵ has reviewed some of their results and it is clear that this group has broken away from the view that intelligence is 'what the tests test for' and is trying to fractionate it into better defined sub-skills, although it is fair to say that no simple alternative definition has emerged. The divergence of opinions becomes even wider if one looks at what sociologists and biologists are writing on the subject, and a brief discussion of three recent books will bring this out.

D.L. Eckberg in *Intelligence and Race*⁶ presents a vigorous and sustained attack on the whole intelligence testing movement, based primarily on the circular nature of the validating procedure pointed out above. It is one of several such books, and although his arguments are not always convincing they certainly undermine one's confidence in many of the conclusions that have been drawn from intelligence tests. What is particularly disconcerting is that these arguments are not answered in the next book, *A Model for Intelligence*, a collection of essays edited by H.J. Eysenck⁷; it was probably felt that they had been adequately answered elsewhere (for example, ref.8), but to an outsider it is strange that Eckberg is nowhere even mentioned. The third book, *Brain and Intelligence in Vertebrates*, is an admirable review by E.M. Macphail⁹, which deals very little with intelligence testing in humans, but all the same it is surprising that he too omits mention of Eckberg, and that, of the first 100 authors' names in his index, not a single one appears in the index of Eysenck's book. All these authors claim to be writing about intelligence, and it is understandable that the subject is in confusion if authors ignore one another's criticisms and scarcely overlap in the work they cite. The underlying problem is surely the absence of an agreed definition of their topic.

Genes, biophysics, or language?

To reinforce the need for reassessment these books contain new hypotheses, or perhaps they are reaffirmations of old ideas, that are interesting but totally unrelated to one another. Eckberg is unconvinced by the evidence for a strongly inherited general factor in intelligence, whereas all Eysenck's contributors accept

this without hesitation. This old controversy will be ignored for the moment because it seems most unlikely that intelligence, almost alone amongst the characteristics of living things, should be determined solely by environmental forces, although we shall see later that education and language may indeed make intelligence unusual in this regard. The new theme in Eysenck's book is that the general factor in intelligence may be based on simple biophysical properties of an individual's brain. The reader must refer to the book for details, but the results reported by Jensen, and others reported by Brand and Deary, show correlations up to about 0.6 between measures of central latency for various tasks and classical measures of intelligence. Is intelligence after all simply a matter of mental speed?

Two other chapters in Eysenck's book, by A.E. and D.E. Hendricksen, look at various measures of cortical evoked potentials, and they too find high correlations with results of classical intelligence tests. It is extremely interesting that relatively simple physiological processes should be advanced as the factors underlying general intelligence and its variability between individuals, but surely what we are talking about is a very special kind of skill having something to do with the way information is handled; even if the physiological mechanisms underlying its variability were as simple as those proposed here, this would not answer the question of what intelligence does for us.

Macphail's conclusions are equally startling. He is completely unconvinced by the classical view that primates and man owe their intellectual pre-eminence to expansion of the neocortex, and especially the association areas. Instead his review led him to two hypotheses, which he states as follows:- "first, that there are no differences, either quantitative or qualitative, among the mechanisms of intelligence of non-human vertebrates; second, that man's intellect is distinguished from that of non-human vertebrates by his possession of the capacity for language". This is as far as one could wish from the Hendricksens' premature attempt to explain intelligence and its variations in terms of the biophysics of conduction along fibres and across

*Kenneth Craik Laboratory, Department of Physiology, University of Cambridge, Cambridge CB2 3EG, UK.

synapses, and it clearly harks back to a more conservative view of the origin of mankind's pre-eminence.

The three divergent suggestions — that intelligence is not inherited (Eckberg), that its variations depend upon simple biophysical factors (Eysenck), and that it rests on the capacity for language (Macphail) — form convenient points for discussion and they bring out once more the importance of finding an answer to the question that has been unanswered for so long that nobody likes to ask it any more.

What is intelligence?

Nobody believes that intelligence is a simple skill that can be described in a few words, but it is nonetheless reasonable to hope for a simple definition of the task that intelligence performs. A fresh attempt to do so is justified by the fact that two relevant theoretical subjects have developed since the lengthy discussions of the topic by the pioneers of mental testing, namely statistical decision theory, and information theory; both suggest succinct definitions of intelligence that complement each other, and they accord well with common sense. The first is "the ability to guess right", the second "the ability to discover unexpected orderliness". They complement each other because discovering order is an important way of improving one's guesswork, and they accord with common sense because intelligence manifests itself chiefly by leading to the right instead of the wrong answer. Note also that the tasks pointed to are not simple ones; intelligence has not been insulted by finding simple definitions of what it does.

Now jumping to the right conclusion, or guessing right, might be thought to imply incautious rashness, so one needs to be reminded that all decisions are fallible simply because the evidence is never complete. Good guesswork requires efficient use of all the available information, and this is the problem that statistical decision theory illuminates by enabling variations of rashness and caution to be distinguished from variations in the amount of the available evidence that is used.

The efficiency of guesswork

Consider first making a simple decision about whether a sensory stimulus is present or not. The decisions involved in the exercise of intelligence of course involve more complex data, but the essence of the problem is clearer in the simple case. A judgement of threshold is now generally stated as a statistical problem: it is supposed that the signal, which is the stimulus the subject is trying to detect, is inevitably contaminated by noise, which is either received with the stimulus or added to it in the sense organ or brain. The subject's task is then to decide whether the message he has just received belongs to the population of messages *signal + noise*, in which case he responds positively, or to the population *noise alone*, in which case he says "no". Two im-

portant developments have come from treating thresholds in this way; it enables one first to understand the effect of the subject's threshold criterion, and second to measure the absolute efficiency of the mechanisms performing the judgement.

If a subject has a low criterion he will respond to weak signals but will often say "yes" when there was no stimulus and he was giving a false positive response to a message from the noise alone population. Such a subject may appear to be more sensitive than one who has a high criterion and very rarely gives false alarms, but the techniques of signal detection theory allow one to measure sensory performance in ways that are not influenced by variations of criterion (see for example ref.10). The importance of this for intelligence is fairly obvious: one who bubbles over with new ideas that are often erroneous differs much from one who cautiously waits until he is certain before revealing the answer, but they do not necessarily differ in intelligence. Thus the first result of statistical decision theory is to show conceptually how the quality of decision making can be separated from the other variable element, rashness or caution, that is also involved in guesswork.

The second result is even more important, for it shows us how to express this quality on an absolute scale. The essential point is that there is a limit to how well a statistical judgement can be performed. If the physical signals are known, the best possible performance can be calculated and a hypothetical ideal observer, who uses all the available information, achieves this performance. A real observer performs worse than this to an extent that would be explained if he only used a proportion of the sample of evidence available to him. This measure of efficiency was proposed by Fisher¹¹ for comparing the merits of statistical tests, was used by Rose¹² to compare the performance of the eye and television cameras on an absolute scale, and was first applied to psychological tests by Hecht *et al.*¹³. Both Rose and Hecht *et al.* showed that human subjects receiving an average of 100 quanta give responses suggesting that they make use of only about 6 of them. The same approach can be applied to more complicated perceptual tasks; it can for instance be shown that a subject offered 100 dots in approximately symmetrical positions makes judgements of symmetry indicating that he only makes use of 25 of them¹⁴. What happens to the 94 unused quanta, or the 75 unused dots, are obviously interesting physiological questions, but the important point here is that the figures for efficiency, 6% and 25%, are hard and absolute in a sense that is new to psychology.

A vast effort has gone into the standardization and validation of intelligence tests, but they remain crude measures because they have no externally defined scale on which the results can be expressed. The one used is derived from the spread of

the test scores obtained on a population: it is as if lengths were defined, not by the average size of a human foot (which would be bad enough in this day and age), but by the standard deviation of these lengths as determined by a somewhat variable instrument. What the use of statistical efficiencies offers is a measure more nearly comparable with those used by physiologists of the last century studying the balances for energy, nitrogen and so forth; it was this work that finally rejected vitalism and put the study of body metabolism on a scientific basis.

Perhaps we can do the same for perception and cognition, but first we must choose more appropriate tasks than threshold judgements. The trouble is that keen eyesight, wider knowledge, or a better memory can improve guesswork as much as intelligence can, so we need to understand theoretically a task more relevant to intelligence than a simple sensory threshold.

Efficiency of learning

The obvious ones to consider are learning tasks, such as Macphail used to compare the intelligence of animals. These require the nervous system to detect that there is a statistical association between two sensory stimuli in the case of conditioning, or between a sensory stimulus and the initiation of a movement in the case of instrumental learning. Now there are limits to what a statistician can tell you about the presence or absence of an association, just as there are limits to what he can tell you about a message belonging to the signal + noise category or signal alone category; this means that one could in principle calculate what an ideal associator could do in a learning situation. The analogue of the observer with a low criterion would be an associator that learned after a very small number of trials, and just as the low criterion observer is liable to give frequent false alarms or erroneous detections, so the low criterion associator would learn rapidly, but would be liable to salivate or peck in response to accidentally associated stimuli. And just as a low threshold is not a proof of efficient sensory detection, so the rapid formation of associations is not by itself a proof of efficient learning. Learning an association is obviously a statistical problem and the efficiency with which the evidence is utilized should surely be measured.

The argument so far has shown that good guesswork is an ingredient of intelligence, that the ability to use incomplete evidence can be measured on an absolute scale, and that this scale could in principle be applied to learning tasks of the type that Macphail used in his comparative study of intelligence. Thus one should be able to answer unambiguously the question whether a fish can learn as efficiently as a human, but be warned that, on the scale so far proposed, it might well be able to do so, at least on tasks chosen to favour the fish! Intelligence must, however, imply more

than efficient learning or good guesswork in specific circumstances: an intelligent person will also guess right in a wide variety of circumstances, including completely new ones. In other words intelligence implies versatile and creative guessing as well as correct and efficient guessing. This is easily said, but the associative task required for versatile guessing is horribly difficult and complicated.

The combinatorial jungle

If one says that certain events are associated one means that they occur together more often than would be expected from the overall frequency of their occurrence. The basic problem is that the number of possible associations rises nearly as the square of the number of possible events for paired associations, nearly as the cube for triplets, and so on. As Boole showed more than a century ago¹⁵, the concepts we habitually deal with correspond to logical functions that are more than simple conjunctions of two or three events, and the possible number of these logical functions rises much more rapidly. This means that the task of looking for associations among the enormous number of signals that are used in a vertebrate brain is that of finding a path through a combinatorial jungle of quite unimaginable complexity. As soon as one glimpses this jungle one realizes that finding useful paths through it must be what intelligence is all about: intelligence is not just a matter of detecting this association or that association, but requires knowledge of the associative structure of a large body of information, and that is a formidable task. Ideas from information theory and the study of language may give more insight on this problem than the pioneers of mental testing could achieve.

Redundancy reduction

Consider first an idea from information theory^{16,17}. If messages are correlated with each other, redundancy is present, and it is in principle possible to find a code that will remove it. Such a code would be adapted to the structure of associations among the messages, so its specific form would store information about this structure. Furthermore if such a code were found and used, messages would become simpler, both because fewer symbols would be required to carry the same information, and because some of the associational structure originally present would have been removed. This could be a way of beginning to clear a path through the combinatorial jungle, and it is interesting to note that some items of intelligence tests seem aimed at detecting the ability to devise such redundancy reducing codes: for instance "Give the next member of the series 1, 5, 10, 16, 23, . . .". This requires identifying the principle upon which the sequence was generated, and once this has been done the numbers (except the first two) become wholly redundant. Such a question may

only determine whether an individual has been well enough trained to be able to recognize certain number-patterns, whereas the postulated redundancy reduction must mainly occur at precognitive levels in the combinatorial jungle. One can, however, imagine tests in which a computer introduces regularity into an otherwise random display and the subject's task is to identify it; by these means one might obtain absolute measures of the mental capacity to detect orderliness, and this is certainly one task performed by intelligence^{18,19}.

With this thought in mind it is amusing to speculate on the possible future of the electronic games machines in an amusement arcade. First, it seems that they may tap and exercise the very skills which, it is here claimed, lie at the root of intelligence. But what is even more exciting is the possibility that they could award scores for their players that would constitute absolute measures showing how well they had utilized the information that was available to them. This would, of course, be the optimal feed back, and if intelligence is trainable the truants in the arcade may soon outshine the conscientious students who remain in school.

Two roles of language

But there is more to intelligence than guessing right at a games machine, so let us consider Macphail's interesting suggestion that it is mankind's capacity for language, and that capacity alone, that marks his brain out from those of the remainder of vertebrates. Now language is often defined solely by its communicative role — it enables one individual to convey his thoughts to another and thereby influence his behaviour. But it has another role, that of organizing the representation of information entirely within one individual's mind²⁰⁻²². This representational role is unfamiliar to many, but it is not necessarily an unimportant secondary aspect of language; quite to the contrary, it could well be the case that the representational role is primary in a developmental and evolutionary sense. On this view mankind's intellectual breakthrough would have been achieved in two stages: the evolution, first of the representational scheme that underlies language, and only after that the development of the ability to communicate with other humans by messages based upon this scheme.

If Macphail is right one must ask which of the two roles of language is most important in giving mankind its superior intelligence; if it were the representational role, and if the argument of this article is correct, then the representational scheme underlying language must be what enables us to find useful paths through the combinatorial jungle. This is an attractive thought and points to an aspect of language that has received little attention²³. But the two roles are not of course independent; because of the communicative role an individual's representational scheme is not just the product of his inherited genes and a passive environment acting through his senses, but also results from the active intervention of the language community to which he belongs. It is perhaps the realization of this crucial role that underlies some people's desire to minimize the importance of the inherited component, and although this may be misguided one can see that intelligence does indeed differ from other biological characteristics in the manner of its dependence on genes and culture.

Conclusions

We saw at the start that intelligence is not written or thought about in a coherent way; it means different things to different people, and even among the community of intelligence testers there are important disagreements. A new look at the problem suggests a simple role for intelligence: it is the capacity to guess right by discovering new order. This not only agrees with commonsense and explains why intelligence is so important for the survival of both man and beast, but also enables us to apply recent knowledge to the problem. Statistical decision theory points to statistical efficiency as an absolute measure for that part of intelligence which is concerned with drawing reliable conclusions from incomplete evidence, and information theory points to the importance of redundancy reduction in handling the overload of information from our sense organs. But this is only one part of the problem: intelligence is concerned with the associative structure of the messages in our brains and this is where language may play an important role.

Decision and information theory have helped us to understand what intelligence does: the next problem is to find out how the task is achieved. □

1. Boring, E.G. *New Republic* 34, 35-37 (1923).

2. Heim, A. *The Appraisal of Intelligence* (Methuen, London, 1954).

3. Tunstall, O., Gudjonsson, G. & Eysenck, H.J. *Bull. Br. psychol. Soc.* 35, 324-331; 333 (1982).

4. Heim, A. *Bull. Br. psychol. Soc.* 35, 332-333; 355 (1982).

5. Hunt, E. *Science* 219, 141-146 (1983).

6. Eckberg, D.L. *Intelligence and Race* (Praeger, New York, 1979).

7. Eysenck, H.J. *A Model for Intelligence*.

8. Jensen, A.R. *Bias in Mental Testing* (Methuen, London, 1980).

9. Macphail, E.M. *Brain and Intelligence in Vertebrates* (Clarendon, Oxford, 1982).

10. Green, D.M. & Swets, J.A. *Signal Detection Theory and Psychophysics* (Wiley, New York, 1966).

11. Fisher, R.A. *Statistical Methods for Research Workers* (Oliver and Boyd, Edinburgh, 1925).

12. Rose, A. *Proc. Inst. Radio Engrs* 30, 293-300 (1942).

13. Hecht, S., Shlaer, S. & Pirenne, M.H. *J. gen. Physiol.* 25, 293-300 (1942).

14. Barlow, H.B. & Reeves, B.C. *Vision Res.* 19, 783-793 (1979).

15. Boole, G. *An Investigation of the Laws of Thought* (1854; reprinted by Dover, New York).

16. Attneave, F. *Psychol. Rev.* 61, 183-193.

17. Barlow, H.B. in *Sensory Communication* (ed. Rosenblith, W.) 217-234 (M.I.T. Press, 1961).

18. Fatmi, H.A. & Young, R.W. *Nature* 228, 97 (1970).

19. Barlow, H.B. *Nature* 228, 1008 (1970).

20. Fodor, J.A. *The Language of Thought* (Cromwell, New York, 1975).

21. McIlver, B. *Proc. 6th Joint Conf. Artificial Intelligence* 1127-1130 (1979).

22. Sloman, A. *Proc. Aslib/BCS Conf.* (Aslib, London, 1979).

23. Barlow, H.B. *Perception* 3, 123-134 (1974).

Moving into the Higgs sector?

Last week's recommendation by US high-energy physicists that the next big accelerator should be a 20 TeV proton-antiproton collider was bold but correct. For there is a new physics to be found.

IN 1913, when Niels Bohr proposed, contrary to Maxwell's well-tested theory of electric charges, that electrons in orbit around a nucleus do not radiate and exist only in discrete orbits, he could scarcely have appreciated the far-reaching implications of these *ad hoc* postulates. His motivation in advancing these ideas was to explain in terms of the structure of the hydrogen atom the spectral emission data that had been accumulated with increasing puzzlement over the preceding decades; and although his postulates succeeded in this task to a remarkably accurate degree, they were from a theoretical point of view deeply dissatisfying. For 13 years, efforts to understand, or even substantiate on more basic grounds, these peculiar postulates were unfruitful. Yet Bohr's boldness in formulating what would now be called fudge factors was vindicated: by formalizing the unexplained, Bohr set up the challenge that was eventually met when the entirely new physics of quantum and wave mechanics was expounded. An explanation of why electron energies are quantized in the hydrogen atom was almost the least of the payoffs.

Today, particle physics is again poised on a springboard of theoretical ad-hockery. And the decision last week by the US High Energy Physics Advisory Panel (HEPAP) to recommend construction of a 20 TeV proton collider is a bold bid to spring off into another new physics, that of the so-called Higgs sector.

Like Bohr's postulates that there are quantized energy states and that electrons which occupy them do not lose energy by radiation, the Higgs mechanism is something of a fudge factor. It is invoked to explain a peculiarity in the Weinberg-Salam-Glashow unification of the electromagnetic and weak interactions; the photon, the carrier of the electromagnetic force, is massless, while the three carriers of the weak force (the W^+ and W^- particles and the neutral Z^0 — the so-called intermediate vector bosons) are very heavy with masses equivalent to almost 100 GeV. The Higgs mechanism is an interaction, perhaps involving an actual particle (the Higgs boson), that accounts for this breakdown of symmetry.

With the discovery of the W and Z^0 particles at the European Centre for Nuclear Research (CERN) in recent months and the stunning success of the electroweak theory in predicting not only their existence but their masses, the excitement has intensified

around efforts to tie up the loose ends of the theory — in particular, the *deus ex machina* of the Higgs mechanism. So loose is this loose end that it is not even clear whether the Higgs mechanism is a particle or a "mathematical fiction" — a fudge factor that incorporates an unknown physics much as Bohr's hydrogen atom represented the then-unknown realm of quantum mechanics by fudge factors.

Whatever the Higgs mechanism may turn out to be, it can on very general theoretical principles be shown to become apparent in the area of 1 TeV; hence HEPAP's motivation in recommending a full commitment of available resources to the construction of a "Superconducting Super Collider" (SSC), also known as the "desert-tron" from an early suggestion that it could economically be built using relatively low-field magnets in a very large ring that would fit nowhere but the southwestern desert. HEPAP's report called the search for the Higgs particle (or its equivalent) the highest priority in particle physics.

The boldness of this decision was all the more remarkable for its accompanying recommendation — essentially without precedent — to abandon a large accelerator project in progress, the Brookhaven National Laboratory's ISABELLE (renamed the Colliding Beam Accelerator, or CBA). That decision was logically necessary, but politically difficult. But to have gone ahead full speed with SSC — an undertaking that will cost \$2,000 million for construction, plus perhaps a similar amount for detectors — in the face of the competition for funds and technical expertise that CBA would pose would have been intolerable. The committee has rightly set its face against such pusillanimity.

The physics that CBA could have carried out did not help its case. Delays in CBA's construction let it be overtaken by the events at CERN. The higher luminosity of CBA's 400 GeV beams would, to be sure, have permitted some experiments that Fermilab's recently completed Tevatron I (1 TeV beams, but lower luminosity) cannot handle. CBA, for example, could have provided a clear count of the number of different neutrinos; it could also have measured within a few parts in 10^4 the W to Z mass ratio. But the Tevatron has other advantages, as do the Stanford Linear Collider (an electron-positron machine, construction beginning this fall) and CERN's large electron-positron ring (also under construction). "CBA is competitive with,

but not significantly better than Tevatron I and other facilities concurrent with it", the HEPAP report found. The argument that the Higgs mass may be much lower than 1 TeV, or that related particles may exist at masses of a few hundred GeV similarly fails to make a persuasive case for CBA. The jackpot lies at 1 TeV, not in the penny-ante games at lower energies. (In order to examine the TeV-level effects, beam energies an order of magnitude greater are needed, whence the design goal of 20 TeV per beam for SSC.)

HEPAP's decision is commendable for its foresight, but also for recognizing the unique convergence in the United States of a physics challenge and a technological possibility. The pioneering work carried out at Fermilab and Brookhaven in the development of superconducting magnets has laid the technical base for the SSC design. The technology has also been helped by CERN's experience with its proton-antiproton collider, which has demonstrated that hard collisions of hadrons at high energies are not as difficult to analyse as had been anticipated. That means it makes sense to go ahead with a proton collider, rather than to try to tackle the much more difficult technology problem of attaining the 1 TeV mass range with an electron machine.

As for Brookhaven, although the loss of CBA is a considerable blow, it would be a tragic mistake for the laboratory to fight HEPAP's decision at the Department of Energy or in Congress. If SSC is to get off the ground, it will need the full support of the physics community; what is more, Brookhaven will surely be needed in research and development work on SSC, particularly in the light of the laboratory's dramatic and successful fight to overcome the problems that cropped up in 1980 in the magnet design for CBA.

Brookhaven may also do well to dust off a once-considered alternative — to turn CBA into a heavy-ion accelerator. The existing tunnel and magnet technology could be used; and although heavy-ion physics may not mesh well with the interests of many physicists now at the laboratory, it is hardly a backwater — there is now intense excitement, virtually a renaissance, in this area of nuclear physics for its potential to yield unique insights into the bulk behaviour of those most elusive bits of the particle-physics world, quarks and gluons. Brookhaven would surely find a receptive ear for such a proposal. **Stephen Budiansky**

Reproductive evolution

Mating types and uniparental transmission of chloroplast genes

from Brian Charlesworth

MANY sexually reproducing species of lower eukaryotes are not differentiated into males and females; all gametes are of equal size. Nevertheless, in a large number of species, gametes can be classified into different mating types, such that successful sexual fusion occurs only between cells of opposite mating type. Such species are called heterothallic, in contrast to homothallic species where no such restrictions on gametic union exist. The common pattern in both algae and lower fungi is to have just two mating types, controlled by alternative mendelian alleles^{1,2}. In higher fungi (Basidiomycetes), many mating types, with more complex inheritance, may be present². Mating types present an interesting problem for the evolutionary theorist. An ancillary problem is posed by the association of mating types with uniparental transmission of chloroplast genes in the green algae *Chlamydomonas*³, described in detail below.

The only serious theoretical study of the evolution of mating types has recently been made by Rolf Hoekstra⁴. He assumes a life cycle characteristic of a yeast or unicellular alga such as *Chlamydomonas*, in which haploid cells proliferate vegetatively by mitotic cell divisions. In appropriate conditions, the vegetative cells can differentiate into gametes, pairs of which can then fuse to form diploid zygotes. The zygotes undergo meiosis to produce tetrads of haploid daughter cells, which resume the cycle of vegetative growth. Hoekstra assumes that the species is initially homothallic, so that gametic union can occur between cells derived by vegetative division from the same meiotic product. Gamete recognition and fusion depends on (at least) two complementary factors, such as a diffusible pheromonal attractant and a corresponding receptor site. A gamete of a homothallic species produces both factors, and might be less efficient at recognition than gametes specialized to produce just one factor, because of self-saturation of the receptor sites. A mutant allele that, say, abolishes production of the pheromone might therefore have a selective advantage and invade the homothallic population, but clearly it cannot spread to fixation. Its presence would facilitate the spread of a mutant at another locus which abolishes production of the receptor sites. Selection would favour tighter linkage between the two loci, leading ultimately to the existence of a supergene involving variation of at least two distinct but tightly linked loci, with two complementary types of gamete predominating in the population. This prediction of a supergene controlling

mating types agrees with genetic data on *Chlamydomonas*⁵, and with both genetic and molecular data on yeasts^{2,6-8}. The model works best if the gamete recognition depends in part on diffusible pheromones; the mating-type loci in yeast are known to control such pheromones^{2,9}.

This model thus provides an attractive explanation for heterothallism, at least for the case of two mating types, but how does one explain the existence of homothallic species? Homothallism would be favoured by selection in environments where population density is low, so that the chance of encounter is small for gametes derived from different vegetative clones. (This is analogous to the selection pressure towards self-fertilization observed in flowering plant species in which pollination efficiency is low¹⁰.) Homothallism in yeast is achieved by an elaborate mechanism in which both sorts of mating-type gene are present within each cell, one of the two being expressed as a result of movement to a special chromosomal site¹¹. In this way, fusions between gametes derived from the same clone are made possible. There is also indirect evidence for the presence of two mating-type genes in the homothallic alga *Chlamydomonas monoica*¹². This system is most easily understood in terms of descent from an ancestor that was originally heterothallic.

As mentioned earlier, mating types in *Chlamydomonas* raise another evolutionary issue. Denoting the mating types in heterothallic species as mt^+ and mt^- , it was discovered by Ruth Sager in 1954 that certain mutant alleles are only transmitted to the progeny of a sexual cross when carried in the mt^+ parent³. It now seems established that such mutants are carried on the chloroplast DNA^{13,14} and that the chloroplast DNA from the mt^- parent is destroyed in the zygote shortly after mating^{15,16}. The chloroplast DNA in the progeny thus comes only from the mt^+ parent and is apparently protected from degradation by methylation of its cytosine residues¹⁶. This is in accordance with an hypothesis of Sager and Ramanis¹⁷. UV irradiation of the mt^+ gametes before mating reduces the frequency of uniparental transmission, showing that the effect of the mt^+ locus is exerted before zygote formation¹⁸.

A new twist to the story has been provided by the work of Van Winkle-Swift and Aubert¹² on the homothallic species *Chlamydomonas monoica*. In crosses between wild type and a mutant conferring resistance to erythromycin (*ery-u1*), they find that tetrads produced at meiosis show

either 100 per cent transmission of *ery-u1* or 100 per cent transmission of wild type, each type of tetrad occurring with roughly equal frequency. They suggest that haploid cells of *C. monoica* contain both mt^+ and mt^- genes, but a gamete expresses only one of these, with fertile unions occurring solely between mt^+ and mt^- . Uniparental inheritance follows the usual mating-type rule, accounting for the two classes of tetrad.

The evolutionary *raison d'être* of this convoluted mechanism of transmission of chloroplast DNA has been somewhat obscure, although Sager¹⁹ has suggested that it has the advantage of suppressing recombination between chloroplast genes. A simple explanation in terms of 'selfish DNA' is as follows. Consider a heterothallic species in which mt^+ and mt^- chloroplast genomes are transmitted in equal frequency to the next generation. A chloroplast mutation arises which interacts with a product of the mating-type region, in such a way that it causes destruction of the mt^- -derived chloroplast DNA when transmitted via an mt^+ gamete. Such a mutation will spread to fixation, since it ensures the exclusion from the progeny of chloroplast genomes derived from the opposite parent. A similar mechanism will operate in homothallic species, if Van Winkle-Swift and Aubert's model is correct and provided that at least some matings take place between gametes derived from different meiotic products. A difficulty is that two functions are required for uniparental transmission: protection of the mt^+ -derived DNA and destruction of unprotected DNA. The postulated variant would thus have to involve a double mutational event. A similar requirement for two mutations exists for systems of unequal transmission of nuclear genes in other organisms²⁰. □

Brian Charlesworth is in the School of Biological Sciences, University of Sussex, Falmer, Brighton, Sussex BN1 9QG.

1. Lewin, R.A. *The Genetics of Algae* (Blackwell, Oxford, 1976).
2. Fincham, J.R.S., Day, P.R. & Radford, A. *Fungal Genetics* 4th edn (Blackwell, Oxford, 1979).
3. Sager, R. *Proc. natn. Acad. Sci. U.S.A.* **40**, 356 (1954).
4. Hoekstra, R. *J. theor. Biol.* **98**, 427 (1982).
5. Gillham, N.W. *Am. Nat.* **103**, 355 (1969).
6. Klar, A.J.S. *et al. Nature* **289**, 239 (1981).
7. Teupold, U. *Cold Spring Harb. Symp. quant. Biol.* **23**, 161 (1958).
8. Meade, J.H. & Gutz, H. *Genetics* **83**, 259 (1976).
9. Wilkinson, I.E. & Pringle, J.R. *Exptl Cell Res.* **89**, 175 (1974).
10. Lloyd, D.G. in *Demographic Factors and Mating Patterns in Angiosperms* (ed. Solbrig, O.T.) (Blackwell, Oxford, 1980).
11. Hicks, J., Strathern, J.N. & Klar, A.J.S. *Nature* **282**, 478 (1979).
12. Van Winkle-Swift, K.P. & Aubert, P. *Nature* **303**, 167 (1983).
13. Spritzer, R.J. & Mets, L.J. *Nature* **285**, 114 (1980).
14. Grant, D.M., Gillham, N.W. & Boynton, J.E. *Proc. natn. Acad. Sci. U.S.A.* **77**, 6067 (1980).
15. Kuroiwa, T. *et al. Nature* **298**, 481 (1982).
16. Burton, W.G., Grabow, C.T. & Sager, R. *Proc. natn. Acad. Sci. U.S.A.* **76**, 1390 (1979).
17. Sager, R. & Ramanis, Z. *Theor. appl. Genet.* **43**, 101 (1973).
18. Sager, R. & Ramanis, Z. *Proc. natn. Acad. Sci. U.S.A.* **58**, 931 (1967).
19. Sager, R. in *Genetics and Biogenesis of Mitochondria and Chloroplasts* (eds Birky, C. W., Perlman, P. S. & Buys, T. J.) 252 (Ohio State University Press, 1975).
20. Charlesworth, B., & Hartl, D.L. *Genetics* **89**, 171 (1978).

Genetics

Sex is simple

from Peter Goodfellow

IN mammals, chromosomal determination of sex is relatively homogeneous: females are homogametic (XX) and males are heterogametic (XY). Consideration of individuals with abnormal sex chromosome contributions has proved that the Y chromosome is male-determining as XO individuals are phenotypically female and XXY individuals are male, although XO and XXY individuals are usually sterile^{1,2}. In other orders of animals different chromosomal sex-determining mechanisms exist (reviewed in ref. 3). Females may be heterogametic as in birds (in female heterogamety the sex chromosomes are known as Z and W, females and ZW and males are ZZ) and sex determination may depend, as in *Drosophila*, on chromosomal dosage rather than the presence of a single sex-determining chromosome. In the latter case XO individuals are male rather than female; in an extreme form of this type of determination the Y chromosome is lost altogether, giving XO males and XX females (O meaning that no second sex chromosome is present in males). The wide variation in chromosomal sex-determining mechanisms, often with differences between members of the same order or even family, poses a problem for the evolution of sex. Assuming that the basic sex-determining mechanism has not evolved independently on several different occasions, how do you progress from one chromosomal sex-determining mechanism to another?

The first part of the answer was provided by Fisher in his classic text, *The Genetical Theory of Natural Selection*⁴. He suggested, for mammals, that a single Y-linked gene controlled the expression of a series of sex-limited autosomal genes. Thus, the complicated male and female phenotypes would be under the control of a single gene. This hypothesis has the advantage that any phenotype could easily fall under sex-control and could explain the wide variations found in sexual dimorphism.

The second part of the answer comes from studying sex in *Caenorhabditis elegans*, a nematode. The male *C. elegans* has an XO constitution, XX animals are

hermaphrodites. In this issue of *Nature*⁵, J. Hodgkin in an elegant series of genetic experiments has created a variant of *C. elegans* which has a ZZ (male), ZW (female) sex-determining mechanism. Remarkably this transformation requires only two mutations, both of which are at the same autosomal locus⁵.

As discussed by Hodgkin, mutations at several loci can affect sex determination; however, the most profound changes are caused by mutations at the *tra-1* locus. Recessive null mutants at the *tra-1* locus convert XX hermaphrodites to phenotypic males which are often fertile⁶. The null mutants have no effect on male XO animals. Several dominant mutants at the *tra-1* locus convert XX and XO animals into fertile females⁷. Mating null *tra-1* (male) homozygotes with dominant *tra-1*/null *tra-1* (female) heterozygotes produces offspring the same as the parental types in a 1:1 ratio⁵. Thus, the experimental nematode has been converted from an

XX-XO dosage system to as ZW-ZZ female heterogametic system.

The important conclusion is that sex is simple: the whole sexual phenotype, both somatic and gonadal, is under the control of the apparently single *tra-1* locus. A similar conclusion could be drawn from recent experiments in mice in which a Y-linked controlling locus, defined by *sxr*⁸, and a series of autosomal modifiers have been described (reviewed in ref. 9). To take the analogy as far as homology it would be interesting to see if the *tra-1* locus is associated with the Bkm sequences known to be linked to the *sxr* locus¹⁰. Of course, the final answer on homology will be obtained from cloning the *tra-1* gene; it can be safely predicted that this sequence will be used to search many mammalian gene banks and will be injected into the zygotes of mice. □

Peter Goodfellow is at the Human Molecular Genetics Laboratory, Imperial Cancer Research Fund, London WC2A 3PX.

Galactic emission lines

Spectra that defy explanation

from C. Martin Gaskell

IN this issue of *Nature* (p.241), A. P. Fairall reports observations which, if valid, challenge the basis for our understanding of an entire category of galaxies. What he has reported are 30-minute variations in the 'forbidden' lines of a southern active galaxy (number 427 in a list of galaxies whose activity has been discovered by Fairall). These variations are about six orders of magnitude too fast to be explained within the framework of our present theories on how the variations are generated, for reasons I shall explain in this article, in which I also want to offer an unspectacular explanation for the findings.

A Seyfert galaxy is a very low-luminosity quasar. Such galaxies are recognized by strong emission lines from photoionized gas clouds in the nucleus of the galaxy. They show two types of line emission: 'narrow' lines whose small Doppler widths imply velocities of a few hundred kilometres per second, and 'broad' lines with Doppler widths of many thousands of kilometres per second. A galaxy is called a Seyfert 1 galaxy if the strong hydrogen lines are dominated by the broad-line component and a Seyfert 2 galaxy if the broad-line component is absent. Fairall 427 is a Seyfert 2 galaxy. Intermediate cases are known as Seyfert 1.5, 1.6, 1.7 and so on. The two components of line emission come from physically distinct regions: the broad-line region and the narrow-line region.

The narrow-line region is dominated by what is called 'forbidden'-line emission. A forbidden line arises when an electron makes a transition that has an extremely low probability (typically 10^8 times lower than a 'permitted' transition), giving rise to

a visible photon. Because the probability of the transition is so low the electron spends a long time in the upper level before spontaneously making the downward radiation transition. If the density of the gas is relatively high it is more likely that the electron will be knocked out of the level to some other level by a passing electron before it has got round to making the forbidden transition that would give rise to a photon. At high gas densities the transition is thus prevented, or 'forbidden'. Whenever we see forbidden lines, therefore, we know that the gas density is lower than some critical value.

This is the case for the narrow-emission-line regions in Seyfert galaxies and quasars which typically have densities of about 10^4 particles per cubic centimetre. The broad-line region, on the other hand, does not show forbidden-line emission so we know that the density is high (in fact about 10^{10} particles cm^{-3}). The emissivity of a gas (number of photons created per unit volume per unit time) is proportional to the square of the density so a much larger volume of gas is needed to produce the narrow lines than is needed to produce broad lines of the same strength. These arguments tell us that the narrow forbidden lines typically come from a region at least 100–1,000 light yr across, while the broad lines come from a region < 1 light yr across.

These size scales are among the best established physical parameters of quasars and quasar-like objects. Photoionization-based estimates of the size of the broad-line region are supported by observations of its variability. Ten years ago A. M. Cherepaschuk and V. M. Lyutyi (*Astrophys.*

- Ohno, S. *Sex Chromosomes and Sex-linked Genes* (Springer, Berlin, 1967).
- Davis, R. J. *med. Genet.* **18**, 161 (1981).
- Austin, C. R., et al. in *Mechanisms of Sex Determination in Animals and Man* (eds Austin, C. R. & Edwards, R. G.) 1 (Academic, London, 1981).
- Fisher, R. A. *The Genetical Theory of Natural Selection* (Clarendon, Oxford, 1930).
- Hodgkin, J. *Nature* **304**, 267 (1981).
- Hodgkin, J. A. & Brenner, S. *Genetics* **86**, 275 (1977).
- Hodgkin, J. *Genetics* **96**, 649 (1980).
- Cattanach, B. M., Pollard, C. E. & Hawkes, S. G. *Cytogenet. Cell Genet.* **10**, 318 (1971).
- Eicher, E. M. in *Prospects for Sexing Mammalian Sperm* (eds Amann, R. P. & Seidel, G. E.) 121 (University Press, Boulder, 1982).
- Singh, I. & Jones, K. W. *Cell* **28**, 205 (1982).

Lett. 13, 165; 1973) discovered that the broad-line region of the Seyfert galaxy NGC4151 varied on a time scale of about two weeks when the photoionizing continuum varied. This suggests that the broad-line region of NGC4151 is a few light-months across. For the narrow-line region the most direct support for the 100–1,000 light yr size comes from direct imaging of nearby active galactic nuclei (to give one a sense of the size of the region it is useful to remember that a typical galaxy, such as our own, is about 100,000 light yr across). A survey of nearby spiral galaxies made by W. C. Keel of Kitt Peak Observatory (*Astrophys. J.* 268, 632; 1983) shows that nuclear gas is usually concentrated in a nearly spherical core with half-intensity diameter of typically 600 light yr, often surrounded by a disc in the plane of the galaxy traceable to over 2,000 light yr diameter.

All of our present understanding of the narrow-line region of Seyfert galaxies as outlined above is quite incompatible with what Fairall reports. The problem is that the forbidden lines Fairall observes must come from a large region perhaps 1,000 light yr across. Such a large region cannot produce variations in half an hour. The shortest time scales we should expect for variability of the forbidden lines he observes is several years (or more likely decades). Our estimate of the size of the region would have to be in error by 10^6 to produce what Fairall claims. Completely unexpected observational results in astronomy have often led to a new understanding of some phenomenon and Fairall reminds us that quasars have surprised us in the past. However in the past, unexpected observational results have quickly produced a variety of plausible physical explanations. Not all of these will ultimately prove correct of course, but at least the explanations exist. In the case of the alleged variability in the line profile in Fairall 427, no conventional physical conditions can explain the observations and the physical conditions are well understood.

What, then, is going on in Fairall 427? My personal bet is that the results are due to the narrow aperture used by the South African observers to take their spectra. It is smaller than typical angular sizes of narrow forbidden-line regions at the redshift of Fairall 427. The slightest error in setting the 1.8 arc s aperture on the nucleus will result in sampling a different part of the forbidden-line region with different velocities giving different line profiles. An error of only 0.25 arc s would probably be enough to produce the observed effect. It will need high-quality spatially resolved spectra taken in good atmospheric conditions with a large telescope that tracks very well to see if this indeed the case. □

C. Martin Gaskell is a lecturer in the Department of Astronomy, University of Texas, Austin, Texas 78712.

Molecular structure

Computer cues to combat hypertension

from David Blow

THE structure of renin, described by Blundell and Sibanda¹ in this issue of *Nature*, should provide a crucial test of the latest technology of molecular drug design. Renin has a specific role in activating the precursor of angiotensin II, an octapeptide hormone which causes contraction of arterial tissue and thus controls blood pressure. Elevated plasma renin activity, the cause of some forms of hypertension, could be controlled if there were a specific renin inhibitor that could prevent activation of the precursor.

Molecular biology, despite its pervasive influence on biological science, has a disappointing record of direct achievement in medicine. There are plenty of new tests and diagnoses, but no cures. Pharmaceutical companies have invested enormously in biotechnology, but most new drugs still come by the classical route of screening obscure natural products, and chemical synthesis of a wide range of analogues to seek improved properties.

The control of angiotensins through a renin inhibitor has already been a subject for 'rational' drug design. Inhibitors have been based on angiotensin precursors² or on pepstatin which is a general inhibitor of the aspartyl proteinases³. These inhibitors were designed without the detailed knowledge of the structure of renin.

Human renin has proved difficult to purify. The proposed structure is based on renin obtained from mouse submaxillary gland and whose catalytic properties are almost identical⁴. The amino acid sequence of this enzyme^{5,6} closely resembles that of pepsin. Apart from a six-residue extension at the amino terminus, renin has seven insertions and three deletions, and the remaining 320 amino acids of mouse renin show 42 per cent identity with porcine pepsin⁶.

Pepsin, the prototype of aspartyl proteinases, was one of the first enzymes to be studied by X-ray diffraction⁷, and its crystal structure is known⁸. More accurate crystal structures are available for highly homologous aspartyl proteinases from the moulds *Endothia parasitica*⁹, *Rhizopus chinensis*¹⁰ and *Penicillium janthinellum*¹¹.

The substrate-binding site of the aspartyl proteinases is formed as a cleft between two lobes of the enzyme molecule. These two lobes, parts of the same polypeptide chain, show considerable similarity of structure. The peptide chain of the substrate is bound between them and subsites for five or six amino acid residues have been identified. At the central part of the binding site two aspartate residues converge on the substrate. These are the residues primarily

involved in the hydrolysis that gives the aspartyl proteases their name.

The binding site of renin is more different from the other acid proteases than, say, *Penicillium* pepsin and porcine pepsin are from each other. These differences may account for the extreme specificity of renin for the angiotensinogen precursor, while the other acid proteinases display a broad-ranging specificity. Blundell and Sibanda¹ point out, however, that the specificity of renin probably depends on other parts of the angiotensinogen molecule remote from the site of activation.

Because all precedents suggest that in such highly conserved structures the conformation of the polypeptide main chain is very little changed, it is probable that the geometry of the specificity site can be predicted rather accurately by analogy. The situation is very reminiscent of the predictions giving a structural explanation of the specificity of trypsin¹² and elastase¹³ which were made when chymotrypsin was the only serine proteinase whose structure was known. The predicted trypsin structure allowed the interaction of trypsin with pancreatic trypsin inhibitor to be accurately predicted¹⁴.

The complexity of protein structure makes accurate three-dimensional analysis of possible interactions extremely difficult. High-performance computer graphic techniques have been developed in the last year or two, and are now taking the place of conventional model building^{15–17}. The striking support for the new Molecular Graphics Society has partly come from pharmaceutical companies. The first issue of the Society's journal has a paper on the

1. Blundell, T.L. & Sibanda, B.L. *Nature* 304, 273 (1983).
2. Boger, J. *et al.* *Nature* 303, 81 (1983).
3. Szelke M. *et al.* *Nature* 299, 555 (1982).
4. Hirose, S. *et al.* *Circ. Res.* 45, 275 (1977).
5. Pathier, J.-J. *et al.* *Nature* 298, 90 (1982).
6. Mosono, K.S., Chang, J.-J. & Inagami, T. *Proc. natn. Acad. Sci. U.S.A.* 79, 4858 (1982).
7. Bernal, J.D. & Crowfoot, D. *Nature* 133, 794 (1934).
8. Andreeva, N.S. *et al.* *J. molec. Biol. (Russ.)* 12, 922 (1978).
9. Blundell, T.L. *et al.* in *Enzyme Regulation and Mechanism of Action* (eds Mildner, P. & Ries, B.) *FEBS* 60, 281 (1979).
10. Subramanian, E. *et al.* *Proc. natn. Acad. Sci. U.S.A.* 74, 559 (1977).
11. James, M.N.G. & Sielecki, A.R. *J. molec. Biol.* 163, 299 (1983).
12. Steitz, T.A., Henderson, R. & Blow, D.M. *J. molec. Biol.* 46, 337 (1969).
13. Hartley, B. *Phil. trans. R. Soc.* B257, 77 (1970).
14. Blow, D.M. *et al.* *J. molec. Biol.* 69, 137 (1972).
15. Humblet, C. & Marshall, G.R. *Drug devl. Res.* 1, 409 (1981).
16. Kuntz, I.D. *et al.* *J. molec. Biol.* 161, 269 (1982).
17. North, A.C.T. *Chemistry and Industry*, 221 (1982).
18. Pearl, L.H. & Honegger, A. *J. molec. Graphics* 1, 9 (1983).
19. Hassal, C.H. *et al.* *FEBS Lett.* 147, 175 (1982).

analysis of molecular surfaces¹⁸. Such tools will now be thoroughly tested in the attempt to use these new results to design a pharmaceutically useful inhibitor of renin.

Meanwhile, Roche has been concentrating on the angiotensin-converting enzyme, which carries out the final step in producing active angiotensin II by hydrolysis of the precursor angiotensin I released by renin. Sophisticated computer graphics were used to survey the likely active confor-

mations of known inhibitors of the converting enzyme¹⁹. This survey guided the synthesis of putative inhibitors with functional groups in rigidly locked orientations. It resulted finally in the synthesis of a potent bicyclic inhibitor molecule, and a patent was applied for a few weeks ago. □

David Blow is Professor of Biophysics at the Imperial College of Science and Technology, London SW7 2BZ.

Immunology

Molecular biology of the major histocompatibility complex

from Janet Lee and John Trowsdale

EARLIER this year geneticists and molecular biologists met for the second time* to discuss how DNA cloning has so far helped to elucidate the structure, function and organization of the genes of the major histocompatibility complex (MHC). The cloning of the class I genes, which was already well advanced at the time of the first meeting last year¹, revealed that at least this family of genes is much more extensive than had been predicted from the serology, and that the exchange of sequences between different members of the family (possibly by gene conversion) may make an important contribution to the extreme polymorphism of these genes. It was at that time unclear whether the class II sequences would also prove more numerous than expected, and the class III genes which encode the apparently unrelated components of complement had not yet been cloned. At the second symposium on the cloning of the MHC, participants reported progress on three fronts: the class II family has, at least in man, proved to be more complex than the serology predicted; the first class III genes have been cloned; and the first experiments on function, using genetically manipulated cloned class I genes to transfect cultured cells, have borne fruit.

According to serology and cellular immunology, the human class II region contains three loci, *DR*, *DC* and *SB*, each encoding two chains, the α and β chains, that assemble into a dimeric molecule. Both cDNA and genomic clones homologous to *HLA-DR* and *-DC* α and β genes have now been isolated, and C. Auffray (Harvard University) and H. Erlich (Cetus) described possible candidate cDNA clones for the *SB* α chain. In an analysis of cosmid sequences, J. Trowsdale (ICRF, London) found, in addition to the *DR*, *DC* and putative *SB* α genes, at least three more possible α genes in the human haploid genome. Although the α chain of the class II genes is generally nonpolymorphic by

comparison with the β chain, the restriction maps of the *DC* α gene and a closely related sequence, *DX*, do show signs of polymorphism; and C. Auffray and J. Silver (Michigan State University) showed that there is indeed a highly variable stretch of residues in the coding sequence for the N-terminal domain of the *DC* α chain. Similar polymorphism is seen in the murine homologue of *DC*, the *I-A* α gene (C. Benoist, Stanford University).

Data reported by P. Peterson (Wallenberg Institute, Uppsala), E. Long and J. Gorski (University of Geneva), and J. Strominger (Harvard University) have led to an estimate of at least six class II β genes. Three *DR* β genes and two *DC* β genes have been identified in cloning studies and by analysis of Southern blots, and possible *SB* β -chain clones were also described. *SB* clones cannot however be conclusively identified until the actual *SB* product has been more fully characterized.

So far, the human class II region seems to be more complex than that of the mouse, in which the number of sequences corresponds more closely to the serological predictions (see ref. 2). However, it remains possible that additional undetected genes are present (R. Flavell, Biogen). Whether the additional class II sequences contribute to polymorphism in class II molecules, as has been suggested for the class I family, thus remains an open question.

Certainly whatever contributes to polymorphism operates in a strikingly uneven way across the MHC. M. Steinmetz (Basel Institute for Immunology) presented evidence from restriction mapping in several mouse strains for two highly polymorphic discrete segments of the *H-2* region. The first extends from the class I *K* genes through the class II *I-A* subregion including the *I-E* β gene, and is followed by a conserved region that includes the *I-E* α gene and the class III genes. The second polymorphic segment surrounds the class I *D* and *L* genes and is followed by the class I *Qa* and *TL* regions, which again are relatively conserved. These data are consistent with serological and biochemical studies

on polymorphism.

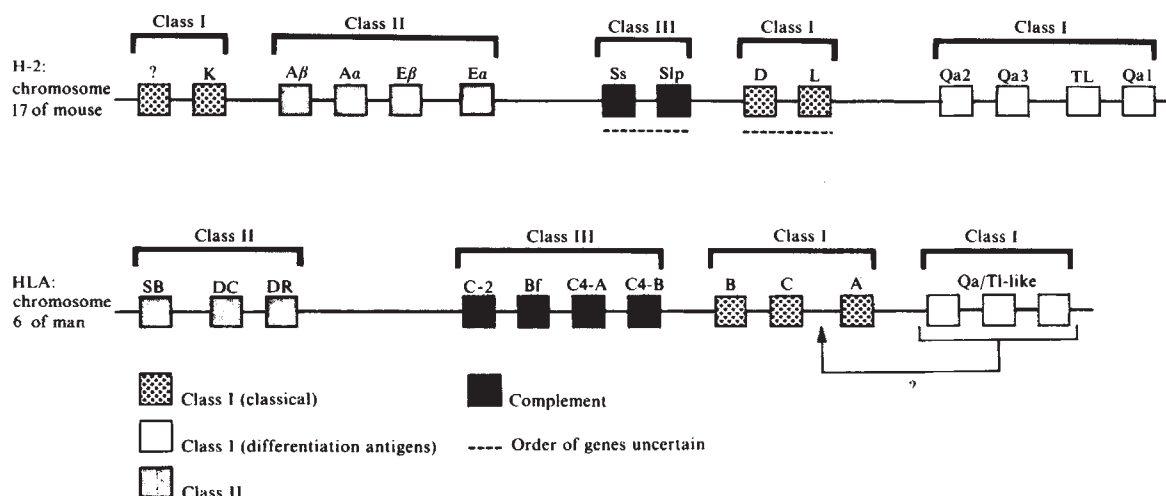
The polymorphism of the *K*, *D* and *L* class I regions, and of the entire class II region, almost certainly reflects selective pressure for variation in the molecules encoded by them. The class I *K*, *D* and *L* molecules function as recognition signals for cytotoxic lymphocytes and the class II molecules function as signals to regulatory lymphocytes; and there is good evidence that their polymorphism increases the range of pathogens against which an effective immune response can be mounted (see refs 1 and 3). The contrasting absence of polymorphism in the adjacent class III and *Qa/TL* regions may be due to equal pressure for conservation. This seems probable in the case of the class III genes, which encode interacting components of a complex cascade; but the function of the *Qa* and *TL* molecules is unknown so it is impossible to guess what selective pressures may bear on them. Nor is it clear whether selective pressure alone can explain the sharp boundaries between polymorphic and nonpolymorphic regions, or whether some special mechanism must be operating to generate discrete areas of polymorphism. A better understanding of the functions of MHC molecules may help to shed some light on the selective pressure acting on their structure, and to this end several laboratories are now using cloned genes that have been genetically altered *in vitro* to transfect mouse L cells.

An innovative way of making mutants to do this sort of experiment was described by P. Kourilsky (Pasteur Institute, Paris). Heteroduplexes of strands from different class I genes were made *in vitro* and inserted into *recA*⁻ *Escherichia coli* hosts in which correction of mismatches is probably required for plasmid replication. Thus a patchwork of sequences from the two parental strands was obtained.

Two groups described preliminary studies with manipulated genes designed to clarify the function of the cytoplasmic domains of the class I antigens. G. Evans (Harvard Medical School) showed that in both alloreactive and influenza- or vesicular stomatitis virus (VSV)-specific systems, recognition by the cytotoxic T lymphocytes is, as expected, based entirely on polymorphism in the two N-terminal extracellular domains. The cytoplasmic domains may however play a part in intracellular signalling, and to investigate this possibility, L cells have been transfected with class I genes whose cytoplasmic domains had been removed; this led however to no observable difference in influenza-restricted killing although killing in the VSV-restricted system seemed less effective. M. Zuniga (Cal Tech, Pasadena) showed similarly that in a lymphocytic choriomeningitis virus-restricted system, deletion of exons 5–8 (all cytoplasmic)

Janet Lee and John Trowsdale are at the Imperial Cancer Research Fund Laboratories, Lincoln's Inn Fields, London WC2A 3PX.

*The second meeting on cloning of the *HLA* and *H-2* regions was held in Airlie, Virginia on 17–19 April 1983.



The major histocompatibility complex (MHC) of mouse and man. The MHC molecules can be divided into three classes on the basis of their structure and function. The class I antigens constitute a single class structurally but fall into two functional groups. The first of these contains the 'classical' class I antigens, first discovered as the transplantation antigens and now known to function as target antigens in the recognition and destruction of virus-infected cells by cytotoxic T lymphocytes. They are expressed on virtually all somatic cells. The second group of class I antigens, loosely defined as lymphocyte differentiation antigens because of their tissue distribution, have no known function. The class II antigens are expressed largely on B lymphocytes and macrophages of the immune system, and are believed to be essential for presenting antigen to the helper and suppressor T cells that regulate the immune response. The class III products are components of the complement system. These maps are based on serological and biochemical data: it is now known that the MHC contains other sequences not detectable serologically.

seemed to have no effect on killing. Clear interpretations of such results are not possible however, because the cytoplasmic domains may be functional only in the effector and not the target cell. It is also possible that the truncated molecules could interact in the membrane with complete class I molecules whose intact cytoplasmic domains would then transmit the necessary effect. Indeed, a general hazard of such experiments seems to be the possibility that a truncated gene will recombine on transfection with a complete homologous gene in the host cell.

Recently, Goodenow *et al.*⁴ showed that homologous recombination of the inserted gene with endogenous genes could occur when very large amounts of cloned class I DNA were used in the transfection experiments: in the Goodenow experiment, this resulted in the expression of a complete gene after transfection with an incomplete one. Recombination between *H-2* genes has now also been reported by D. Margulies (Harvard Medical School), who found recombinants between two different class I genes when mouse L cells were transferred with both together. The possibility of homologous recombination is clearly likely to complicate the interpretation of functional tests.

This year has seen the first success in transfecting mouse cells with human genes; and this poses different problems in interpretation, as noted by F. Lemmonier (Centre d'Immunologie, Marseilles). He showed that an L cell transfected with a probable *HLA-3* gene reacted with an antiserum specific for *HLA-9*, and monoclonal antibodies specific for human β_2 -microglobulin (β_2m) variably reacted with some of the transformants even though the human gene for β_2m was absent. Therefore, the dimer of a human class I

gene and mouse β_2m reacted, in some cases, as if it had an altered specificity, showing again that one must be extremely cautious when attempting to identify human genes by serological experiments using transfectants.

It had been expected that expression of class II genes in transfected cells might be more difficult to achieve, first because, unlike class I genes, they are normally expressed only in particular specialized cells; and second, because it was expected that it might be necessary to supply not only genes for the α and β chains, but also the gene for the so-called invariant chain. The invariant (Ii) chain becomes associated transiently with the α and β chains in the cytoplasm, before the appearance of the class II molecule at the plasma membrane⁵. cDNA clones for the invariant chain have now been isolated in man (E. Long, University of Geneva and V. Quaranta, Scripps) and mouse (Singer, Cologne University). The gene is expressed only in B-cell lines and it is not present on chromosome 6 with the *HLA* region⁶ or on chromosome 17 in the mouse⁷.

Cell-surface expression of *DR* α and β gene products has now been achieved after insertion of the appropriate genes into L cells — though, as it turns out, it is not after all necessary for surface expression (B. Mach, University of Geneva). B. Malissen (Cal Tech) found similarly, while examining not surface expression but protein synthesis, that superinfection of the original *I-A* α and β transformants with the Ii chain did not enhance expression. Those transformants can now be used in functional assays.

The complement, or class III, genes made their debut in the agenda for the meeting this year. cDNA and/or genomic clones have now been isolated using

specific oligonucleotides for C4 and factor B. M. Carroll (University of Oxford) described studies with C4 clones indicating the presence of two and possibly three C4 genes in the haploid genome. Factor B clones were used by D. Woods (Childrens Hospital, Boston) to map more precisely the complement genes by hybridization *in situ* to translocation chromosomes, confirming the current map positions between *DR* and *HLA-B*. No restriction-enzyme-linked polymorphism has yet been observed in or around the factor B gene, even after experiments with 23 enzymes and a large panel of DNAs. This is consistent with Steinmetz's evidence mentioned earlier that the region between *I-A* and *K* in the mouse and including the complement genes is relatively conserved. The factor B gene structure has been characterized by D. Campbell (University of Oxford). Its coding sequence is split into 11 exons, and the residues forming the active site are encoded in three different exons, similar to other serine proteases.

It is hoped that the various subregions of the major histocompatibility systems will soon be linked by chromosome walking, and that further analysis of the DNA clones will shed new light on the organization and structure of the system. It should furthermore be possible now to study in a more controlled way the functions of MHC products in antigen recognition and cell interaction in general; and even perhaps their participation in gene conversion. □

1. Robertson, M. *Nature* **297**, 629 (1982).
2. Steinmetz, M. *et al.* *Nature* **300**, 35 (1982).
3. Matzinger, P. *Nature* **297**, 628 (1982).
4. Goodenow, R. *et al.* *Nature* **301**, 388 (1983).
5. Jones, P. *et al.* *Immunogenetics* **16**, 51 (1978).
6. Long, E.O. *et al.* *Proc. natn. Acad. Sci. U.S.A.* (in the press).
7. Day, C.E. & Jones, P. *Nature* **302**, 157 (1983).

Meteoritics

Iron meteorites and the nature of asteroidal cores

from Edward R.D. Scott

ONE of the more munificent features of the clockwork of the Solar System is the occasional delivery of lumps of extraterrestrial material to the Earth's surface. Many of these are chondrites — silicate meteorites whose composition and petrology betray little alteration since their condensation in the solar nebula, and whose geochemistry can provide unique information on the early evolution of the Solar System. By contrast, iron meteorites, which consist primarily of iron–nickel alloys, seem to have had more eventful lives; they have long been thought to have originated within 'parent bodies' that subsequently disintegrated. Recent laboratory studies of the distribution of elements between solid and liquid iron–nickel support the theory that most iron meteorites are samples from asteroidal cores that fractionally crystallized.

Models based on these studies have reproduced some, but not all, of the systematic chemical variations in several groups of iron meteorites. Unresolved problems include the nature of crystal growth, the extent of liquid trapping and crystal settling, and the apparent lack of sulphur in the parental cores of iron meteorites. The resolution of these problems should help to illuminate processes in planetary cores.

It seems reasonable to attribute large chemical fractionations within groups of iron meteorites to igneous processes because metal in chondrites, which have never been molten, is much more uniform in composition. In the simplest crystallization model, elements are distributed between solid and liquid at their equilibrium levels, and there is complete mixing in the liquid but no diffusion in the solid. According to this fractional crystallization model, samples of solids will define straight lines on logarithmic plots of elemental concentrations. In general, concentrations of elements such as nickel, iridium, gold and phosphorus in iron meteorites do define rather impressive straight lines on such logarithmic plots. However, to test the theory conclusively and to learn more about the crystallization of asteroidal cores, it is necessary to model the process using experimentally determined distribution coefficients.

For many geological examples of silicate melts that have fractionally crystallized, the composition of the original melt is believed to be close to that of material at the edge of the magma chamber, which crystallized rapidly before the magma changed its composition. However, there are no equivalent samples for the iron meteorites, and modellers are free to

choose any cosmochemically reasonable starting composition. Cosmochemists also lack the field constraints available to geochemists, and they know much less about the completeness of their sample. However, the large variations of some elements within iron meteorite groups (for example the 10^2 – 10^3 -fold range in iridium concentrations) suggest that we have reasonably generous samples from several asteroidal cores.

Metallurgists at Lehigh University and geologists at the University of Arizona have measured the equilibrium distribution of many elements between solid and liquid iron, and have determined how minor elements affect these distributions. J. Willis and J.I. Goldstein¹ studied the effect of phosphorus, sulphur and carbon on the distribution of six trace elements, and successfully modelled the behaviour of palladium, gold and nickel in four groups of meteorites. C. Narayan and J.I. Goldstein² measured the effect of phosphorus on the distribution of nickel and germanium between solid and liquid iron. The germanium distribution in the largest group of irons (IIIAB) is especially unusual as meteorites define a boomerang-shaped array on a germanium–nickel diagram, instead of plotting linearly. Narayan and Goldstein found that germanium is concentrated in liquid iron when phosphorus concentrations in the liquid are low, whereas at high phosphorus concentrations it prefers the solid phase. Since phosphorus and nickel concentrations in the liquid become progressively higher during crystallization, Narayan and Goldstein were able to explain the change in slope on the germanium–nickel plot from positive to negative for group IIIAB irons. Their successful modelling of this reversal seems to be a convincing proof that iron meteorites are samples of fractionally crystallized melt. However, J.H. Jones and M.J. Drake³, who measured the effects of sulphur on the partitioning of seven elements, could not satisfactorily model the nickel–germanium trends or the distribution of tungsten and chromium in iron meteorites. As a result they are more cautious than the Lehigh group in arguing for fractional crystallization.

Sulphur, which is very soluble in molten iron but almost insoluble in solid iron, should be concentrated in asteroidal cores. However, there are indications from fractional-crystallization modelling that some of the cores that provide us with iron meteorites have lost their full sulphur complement. A. Kracher and J.T. Wasson⁴ of the University of California, Los Angeles

consider the possibility that sulphur was volatilized from the parent asteroids or separated into an immiscible liquid in the core. Their preferred explanation, however, is that the cores contained two separate liquids which formed by episodic melting during core formation. The first liquid to form would be close to the Fe–FeS eutectic in composition (85 wt per cent FeS), and the second would melt at a temperature almost 500K higher and be virtually free of sulphur. Kracher and Wasson argue that there was not time during the million-year lifetime of the liquid core for complete mixing of the two liquids.

Irrespective of whether the parental melts for the iron meteorites lost some sulphur, there should be more iron meteorites rich in iron sulphide that crystallized from the last few per cent of liquid. Only 1 of about 600 iron meteorites and the metallic phase in a single pallasite have compositions close to that of Fe–FeS eutectic; most iron meteorites have only 0.1–1 wt per cent sulphur. What happened to the sulphide-rich iron meteorites? Probably many did not survive the 100–1,000-million-year journey through space as sub-metre bodies because they were fragmented by collisions more easily than sulphide-poor irons^{4,5}. They would also be more susceptible to weathering on the Earth.

Cosmochemical modelling of fractional crystallization provides very little information about the physical distribution of solid and liquid during core solidification. Did solidification begin at the centre, edge or on nuclei within the core, and was the solid liquid interface planar or dendritic? Narayan and Goldstein² argue by extrapolation from their data that there were dendrites about 500 m wide inside the cores, but Jones and Drake³ favour plane-front growth. There is evidence in the texture of stony-iron meteorites that there was substantial and vigorous mixing of molten metal and silicate at the core–mantle boundary. Thus conditions in the core were probably far from equilibrium, and may have been affected by collisions between asteroids⁶.

Direct clues to conditions in asteroidal cores during solidification have been obtained from studies of the structure and chemical variations in the Cape York shower of nine iron meteorites, which

1. Willis, J. & Goldstein, J.I. *J. geophys. Res.* **87** (Suppl.), A435 (1982).
2. Narayan, C. & Goldstein, J.I. *Geochim. cosmochim. Acta* **36**, 259 (1982).
3. Jones, J.H. & Drake, M.J. *Geochim. cosmochim. Acta* (in the press).
4. Kracher, A. & Wasson, J.T. *Geochim. cosmochim. Acta* **46**, 2419 (1982).
5. Goldstein, J.I. & Axon, H.J. *Naturwissenschaften* **60**, 313 (1973).
6. Scott, E.R.D. in *Asteroids* (eds Gehrels, T. & Matthews, M.S.) 892 (University of Arizona Press, 1979).
7. Buchwald, V.F. *Handbook of Iron Meteorites* (University of California Press, 1975).
8. Esbensen, K.H., Buchwald, V.F., Malvin, D.J. & Wasson, J.T. *Geochim. cosmochim. Acta* **46**, 1913 (1982).
9. Esbensen, K.H. & Buchwald, V.F. *Phys. Earth planet. Inter.* **29**, 218 (1982).

weigh between 1 kg and 31 tons⁷. K. Esbensen and colleagues^{8,9} at the Technical University of Denmark found surprisingly large differences in composition among these samples. For example, they found twofold variations in the concentrations of iridium and gold, and up to 10 per cent change in iridium concentration over 0.9 m in a 20-ton specimen. Such variations are not compatible with radial plane-front solidification of a core 10–50 km in radius. Esbensen and colleagues favour dendritic growth or growth from blocks that were detached from the outer edge of the core and accumulated nearer the centre.

Through the application of metallurgical, geochemical and analytical techniques, and the concerted efforts of meteorite hunters and curators across the world, we are beginning to learn much about the nature and formation of metallic cores of asteroids. Such studies, when augmented by field studies of asteroids, should lead to a greater understanding of the more inaccessible planetary cores. □

Edward R.D. Scott is in the Institute of Meteoritics and Department of Geology, University of New Mexico, Albuquerque, New Mexico 87131.

Astronomy

Do we live in an elliptical galaxy?

from M. G. Edmunds

It is perhaps surprising, at a time when a great deal of astronomical research is directed to probing the most distant parts of the observable Universe, that fundamental details of the structure of our own Galaxy should remain uncertain. We know quite a lot about its most conspicuous part — the thin disc of stars, gas and dust in which active star formation is occurring, and which contains our own Sun — but the faint 'spheroid' of old stars which surrounds the disc is less well understood. Two recent, and conflicting, studies of the number of stars in this spheroid highlight the uncertainty; and together with new information on the dynamical properties of the spheroids of other galaxies, they show that we can no longer ignore the possibility that the spheroid may be a major (or even dominant) component of the Galaxy.

Since the ages and distribution of stars in the spheroid may be very like that of an elliptical galaxy, it is possible to speculate that our Galaxy is simply an elliptical which happens to have been born with, or acquired — perhaps very early on — a disc of gas in which classic spiral structure subsequently developed.

The difficulty of judging the importance of the spheroidal component comes from its intrinsic faintness. It contains few (if any) of the short-lived, bright and recently formed massive stars that delineate spiral structure in the disc. Furthermore, although the local density of spheroid mass near the Sun is very much less than that of disc stars, the total numbers of stars could be comparable when summed over the much larger volume occupied by the spheroid.

Stars can now be counted on photographic plates by measuring machines with computer-controlled scanning light beams, with a vast increase in speed and objectivity over earlier manual and ocular methods. G. Gilmore and N. Reid¹ have used these new techniques, linked with fairly extensive calibration of

stellar types and distances from infrared photometry, to push the determination of the relative numbers of stars of different intrinsic brightness to about 100 times fainter than previous reliable determinations. Then, by counting the relative numbers of stars of different observed brightness and colour towards the South Galactic Pole (that is, up through and out of the plane of the Galaxy), they determine the distribution in space of the stars². They identify not only the well known thin galactic disc, but also what they claim is a 'thick' disc extending above the plane some four times further than the thin disc. This thick disc may represent a somewhat flattened spheroid distribution, although (as I shall argue in detail elsewhere) the resultant mass of such a spheroid could be large — indeed it could be between one-third and three times the mass of the disc. Such a thick disc or spheroid (and the two models cannot be distinguished on the basis of the present star counts) would then have to be accepted as a critical component in the formation and evolution of the Galaxy.

Examples of dominating spheroids are known in other spiral galaxies. A recent series of papers by P. C. van der Kruit and I. Searle shows a range from fairly negligible spheroids right up to cases (see, for example, ref. 3) in which the visible spheroid has at least as much mass as the disc, and probably more. It has become customary to distinguish between 'spheroid', which means the visible stars which are not part of the disc, and the 'halo' which means the almost spherical distribution of dark (or almost non-luminous) matter which is invoked to explain the behaviour of rotation velocity as a function of radius in our own and other spiral galaxies. While the rotation curves seem to have forced us to accept a dominance of the halo mass, the contribution of spheroids has, until recently, remained uncertain.

A somewhat different picture from that of Gilmore and Reid has been painted by J. N. Bahcall *et al.*⁴. Their model is

developed by fitting together their interpretation of star count data (which does not include the recent work of Gilmore and Reid) with data on stars near the Sun which show the high-velocity kinematics required for them to spend most of their lives outside the disc, and the kinematic data on the overall rotation of the Galaxy. Adding a concentrated mass component at the centre of the Galaxy, the relative masses they deduce for disc/spheroid/centre/dark halo are 5.6: 0.27: 1.1: 56. In other words, a dominant dark halo, but a visible spheroid which is rather negligible compared with the disc. Gilmore and Reid can say nothing about the dark halo (except that its mass is not in the form of nuclear-energy-generating stars), but the reason for the disagreement on the importance of the visible halo is not obvious. It may lie in the different relationships between intrinsic brightness and numbers assumed for the spheroid stars by the two groups, but the true picture may only surface in later debate.

Further interesting evidence of the relationship between spiral and elliptical galaxies has come from the comparison of the kinematics of stars in the spheroidal components of spirals with the kinematics of stars in ellipticals. A few years ago much excitement was generated by the realization that bright massive elliptical galaxies were not flattened into their elliptical shape by rotational forces. The result was a flurry of theoretical work on the dynamics of such systems, and it was subsequently observationally demonstrated that, in contrast, the spheroidal components of spiral galaxies did rotate fast enough to provide their observed flattenings. This argued for a fundamental difference between ellipticals and spheroidal components. But the spheroids of spirals are typically considerably less luminous than the large elliptical galaxies that had been observed, and a recent paper by Davies *et al.*⁵ has shown that ellipticals of comparable luminosity to the spheroids of spirals do rotate in the same way. Thus it is the bright massive ellipticals which are in some way unusual, and the spheroids of spirals appear very similar to the ellipticals of comparable brightness.

The true nature of the dark halo matter in our own and other galaxies remains an embarrassing problem, but the recent research indicates that the study of what we can see — the spheroid stars in our Galaxy — may be more rewarding than has often been realized. The spheroid may be the closest elliptical galaxy that we can study. □

M. G. Edmunds is at the Department of Applied Mathematics and Astronomy, University College, Cardiff CF1 1XL.

1. Gilmore, G. & Reid, N. *Mon. Not. R. astr. Soc.* **201**, 73 (1982).
2. Gilmore, G. & Reid, N. *Mon. Not. R. astr. Soc.* **202**, 1025 (1983).
3. van der Kruit, P.C. & Searle, I. *Astr. Astrophys.* **110**, 79 (1982).
4. Bahcall, J.N., Schmidt, M. & Soneira, R.M. *Astrophys. J.* **265**, 730 (1983).
5. Davies, R.L., Efsthathiou, G., Fall, S.M., Illingworth, G. & Schechter, P.L. *Astrophys. J.* **266**, 41 (1983).

Infrared astronomy

IRAS circular 2

from the IRAS working group

THE Infrared Astronomy Satellite (IRAS) continues to work well and the IRAS science team presents here the second circular of new far-infrared sources discovered by IRAS. The main scientific goal of the mission is to carry out an all-sky survey in the wavelength range 10–100 μm and the IRAS catalogue is expected to be published during the second half of 1984. However because of the intense interest in the new IRAS sources from infrared and other astronomers, the science team has promised to publish regular circulars containing selected new sources. The team also hopes that observations of these sources at infrared and other wavelengths will provide useful feedback for the planning of the remainder of the mission.

The first circular (*Nature* 303, 480; 1983) was the product of an intensive effort by science team members to identify IRAS sources with interesting classes of astronomical object. This proved to be a very time-consuming exercise and as most science team members are engaged either in an operational role or in data analysis the circulars would have come out rather infrequently had this format been maintained.

Instead it has been decided that in future

the bulk of the sources for the circulars will be chosen at random from a subset of the IRAS data base, with certain restrictions (see below).

IRAS is a joint project between the USA, the Netherlands and the United Kingdom. □

from H. Habing and G. Neugebauer

STARTING with this issue, each IRAS circular will contain about 30–40 objects, chosen either at random from a subset of the IRAS database of objects already processed or by some member of the IRAS Science Team because the sources have particular interest for non-IRAS observations.

The randomly selected objects will be chosen from an area 20° wide in ecliptic longitude that has been recently processed during preparation of the IRAS catalogue of point sources. About half of them will be within 25° of the galactic plane. Further selection criteria are: (1) the source must have a signal-to-noise ratio of >30 in at least one wavelength band if within 25° of the galactic plane and >15 if outside; (2) the source must have been observed on at least 4 orbits; and (3) the source must have a colour temperature between 60 and 12 μm or 25 and 12 μm of less than 350K.

We shall thus be presenting a relatively unbiased microsample of the IRAS survey catalogue with normal stellar sources removed. We hope that by following these procedures we shall ensure a bi-weekly issue of the circulars. □



Photograph of a false-colour image of the IRAS 100- μm data in the vicinity of IRAS 0344 + 327P01, shown arrowed, believed to be a newly forming star still completely shrouded in the cloud of gas and dust from which it is forming. The 'protostar' lies in a region of the sky identified at optical wavelengths as the dark cloud, Barnard 5, thought to be about 1,000 light years away. The white areas are regions of high 100- μm brightness. The temperature of the dust around the protostar, about 180K, high for interstellar material, suggests that the star is almost fully formed, and its low luminosity shows that its mass is probably no greater than that of the Sun. The white area in the lower portion of the photograph shows intense emission from hotter dust (300K) within a cloud of ionized gas (an 'HII region'), excited by far more massive young stars.

Source name IRAS	Coordinates RA/dec	12 μm	25 μm	60 μm	100 μm	Source name IRAS	Coordinates RA/dec	12 μm	25 μm	60 μm	100 μm
0225 + 725P02	02 h 25 min 02 s +72° 30.6'	0.67	0.91	4.3	9.5	0413 + 122P02	04 h 13 min 47 s +12° 17.6'	<0.3	<0.3	2.2	3.2
0225 + 727P02	02 25 50 +72 46.1	1.1	1.8	7.9	35	0414 + 014P02	04 14 57 +01 24.9	<0.3	0.51	2.3	<1
0253 + 604P02	02 53 13 +60 27.8	1.2	11	<1	<6	0422 + 097P02	04 22 39 +09 44.6	<0.4	0.45	1.7	3.7
0254 + 605P02	02 54 54 +60 32.0	0.93	1.6	14	87	0425 + 106P02	04 25 06 +10 37.4	<0.2	0.48	1.7	5.2
0257 + 700P02	02 57 13 +70 02.6	0.55	0.94	5.9	9.6	0426 - 038P02	04 26 17 -03 52.7	0.24	<0.2	1.4	3.6
0259 + 601P02	02 59 53 +60 08.5	0.86	3.8	25	<8	0428 + 075P02	04 28 29 +07 31.4	0.27	0.562	3.1	6.8
0305 + 596P02	03 05 46 +59 41.4	1.7	1.5	34	100	0429 + 066P02	04 29 18 +06 40.2	0.24	0.32	1.8	5.2
0307 + 607P02	03 07 52 +60 46.0	12	16	5.0	<5	0429 - 058P02	04 29 25 -05 51.8	0.34	0.28	2.9	5.3
0313 + 599P02	03 13 31 +59 58.9	1.8	2.1	5.3	27	0433 - 032P02	04 33 36 -03 15.0	<0.2	0.21	1.6	6.1
0314 + 601P02	03 13 31 +60 11.3	1.2	1.3	30	60	0434 - 002P02	04 34 04 -00 14.8	0.28	0.33	1.4	7.1
0318 + 633P02	03 18 12 +63 21.0	0.67	0.57	5.5	18	0437 - 049P02	04 37 45 -04 57.8	<0.3	0.27	1.5	4.2
0320 + 613P02	03 20 03 +61 21.6	0.35	3.1	<0.5	48	0440 + 005P02	04 40 21 +00 31.5	<0.2	0.52	1.9	7.0
0326 + 710P02	03 26 38 +71 02.6	2.2	4.8	2.5	<3	0446 - 049P02	04 46 07 -04 54.4	<0.2	0.38	2.5	2.8
0341 + 678P02	03 41 45 +67 51.6	<0.2	0.59	4.1	28	0447 - 024P02	04 47 28 -02 28.5	<0.2	0.27	3.1	4.5
0353 + 625P02	03 53 44 +62 35.8	2.2	3.8	6.1	8.6	0448 - 055P02	04 48 16 -05 30.2	<0.2	<0.3	1.0	3.8
0402 + 696P02	04 02 35 +69 40.7	<0.2	<0.2	4.2	29	0449 - 063P02	04 49 14 -06 18.9	<0.2	0.38	2.8	8.5
0412 + 085P02	04 12 59 +08 32.8	<0.2	<0.3	2.2	7.4	0450 - 044P02	04 50 49 -04 27.0	<0.4	0.23	1.5	3.5
0413 + 702P02	04 13 47 +70 16.1	0.62	2.3	1.5	<2	0457 - 034P02	04 57 45 -03 25.5	<0.2	0.36	1.9	5.7
						0502 - 043P02	05 02 18 -04 21.8	<0.2	<0.2	1.1	15

The source name consists of four parts: (1) 'IRAS' indicates the origin; (2) right ascension (RA) in hours and minutes, seconds omitted; (3) declination (dec) in decimal degrees, multiplied by 10 and then truncated (i.e. +32° 42.3 = +327); (4) an appendix starting with 'P' and followed by the number of the circular; this appendix stresses that the data are preliminary. The position is given at Equinox 1950.0, and the measurements were made between epochs 1983.1 and 1983.3.

Retroviral transforming genes in normal cells?

Peter H. Duesberg

Department of Molecular Biology and Virus Laboratory, University of California, 229 Stanley Hall, Berkeley, California 94720, USA

The hallmark of retroviral transforming genes (onc genes) are specific sequences which are unrelated to essential virion genes but are closely related to sequences in normal cells. Viral onc genes probably originated from rare transductions of these cellular sequences by retroviruses without onc genes. Consequently, it has been suggested that retroviral transforming genes are present in normal cells in a latent form. However, recent structural analyses indicate that viral onc genes and cellular genes, which share specific sequences, are not isogenic. They differ from each other in scattered point mutations and in unique coding regions. The cellular genes containing onc-related sequences are expressed in normal cells compatible with a normal function. There is as yet no functional or consistent circumstantial evidence that these cellular genes cause cancer in animals that are not infected by viruses with onc genes. Therefore, it is still uncertain whether the onc-related cellular genes have oncogenic potential beyond their role as progenitors of retroviral onc genes.

ANIMAL retroviruses fall into two classes¹, the rare, highly oncogenic class with transforming (*onc* or *onco*) genes and the almost ubiquitous class without *onc* genes. Retroviruses with *onc* genes are unique among viruses because they are not transmitted either in the germ line or by infection. Indeed, they are only known because of their isolation from rare, spontaneous tumours. No epidemics of Rous sarcoma virus (RSV), avian myelocytomatosis (MC29) virus or the Kirsten, Harvey or Moloney murine sarcoma viruses have ever been reported²⁻⁴. Therefore, it was proposed in 1969 that transforming genes of retroviruses pre-exist in normal cells as latent cancer genes that may be transduced by retroviruses without *onc* genes or activated by carcinogens⁵. However, it would appear paradoxical for normal cells to have evolved and maintain a battery of perhaps 16 known oncogenes which, when carried by retroviruses, are the fastest acting and most inevitably carcinogenic agents known to date. The *onc* genes of retroviruses function as direct initiation and maintenance genes of transformation, because all infected cells become transformed within a few days post-infection, and because the transformed phenotype of cells infected by viruses with temperature-sensitive *onc* genes is temperature sensitive^{2,3}.

A comparison of the structure of *onc* genes in retroviruses with their sequence counterparts in the cellular genome has provided the basis for a biochemical solution of that paradox^{1,6}. As proposed, closely-related counterparts of *onc* sequences exist in normal cells^{3,6}. In rare instances, retroviruses have transduced these sequences from cells to generate viral *onc* genes. It is on this basis that these cellular sequences are widely suspected to be latent cancer genes although there is as yet no conclusive evidence for functional equivalence between viral *onc* genes and cellular prototypes. Indeed, recent analyses indicate that the cellular *onc*-related sequences belong to genes that are structurally different from viral *onc* genes. The cellular genes differ from viral counterparts in scattered point mutations as well as in unique coding regions and are expressed in normal cells, compatible with a normal cellular function.

Oncogene and provirus hypotheses

Early experiments by Gross, Kaplan and Huebner showed that infectious leukaemia and sarcoma viruses appear in tumours of mice, induced either by radiation or chemicals, and led to the suggestion that latent tumour viruses exist in normal cells and could be activated like temperate phages^{2,4}. Huebner and Todaro proposed that unexpressed oncogenes are present in all normal cells and that activation of these genes by retroviruses or carcinogens leads to cancer⁵. This view has since become known as the oncogene hypothesis.

In the light of the discovery of reverse transcriptase, Temin proposed the provirus hypothesis which postulated that normal cells contain potential (proto) oncogenes which by somatic mutation and retroviral or cellular reverse transcription could become viral or cellular oncogenes⁷. The critical distinction of the provirus hypothesis from the oncogene hypothesis was that a qualitative rather than a quantitative change was postulated to convert normal cellular genes to oncogenes. Both hypotheses remained essentially untestable until viral *onc* genes were defined.

Viral *onc* and cellular proto-*onc* genes

During 1970-73, the first viral *onc* gene, the *src* gene of RSV, was biochemically defined by analysis of *src* deletion mutants⁸⁻¹¹ and the genomes of several replication-defective oncogenic retroviruses, the Harvey, Kirsten and Moloney murine sarcoma viruses, were first identified and separated from helper viruses¹². The biochemical hallmark of *src* and all other retroviral *onc* genes are specific sequences unrelated to essential virion genes. Accordingly, *onc* genes are not essential to the virus and may be viewed as parasites. Therefore, there are no structural limitations on *onc* gene sequences except that they fit into a retrovirus. As *onc* genes offer no survival value to their retroviral hosts they are readily deleted. Indeed, spontaneous deletion of *onc* genes has been the very basis for their discovery. The *src* gene, for example, is deleted from within the RSV RNA genome that also contains the three essential variant genes *gag*, *pol* and *env* (Fig. 1b)⁸⁻¹¹. The *onc* genes of all other known retroviruses are part of physically independent but replication-defective RNA genomes (see MC 29; Fig. 1a) that are replicated by non-defective helper viruses. Such parasitic defective-transforming viruses are readily lost from their helper viruses during passage of the helper virus complex in animals or in cell cultures^{13,14}. The virus with the transforming gene would be further counter-selected against by interference with helper viruses. This is a probable reason that epidemics of retroviruses with *onc* genes have never been observed. By contrast, the pathogenic and oncogenic genes of all other RNA and DNA viruses perform survival functions and hence cannot be deleted without affecting the viability of the virus.

About 16 different *onc*-specific sequences have so far been identified in various retroviruses^{1,3,6}. Table 1 lists these *onc* genes and the viral prototypes in which they were first diagnosed, together with a brief description of their oncogenic properties. Some of these *onc* genes regularly cause multiple types of cancers. For example, the *onc* gene of MC29 causes myelocytomatosis, carcinoma and sarcoma, while other *onc* genes like *src* of RSV or the *onc* gene of avian myeloblastosis

virus (AMV) cause one predominant type of tumour. Despite sometimes causing cancer in the same tissues as each other, *onc* genes of different viruses can be structurally unrelated. (The argument that different *onc* genes with overlapping tissue specificity may make functionally identical products fails to explain the unique specificities of some of these *onc* genes.)

The biochemical definition of *onc* genes provided the basis for testing their postulated cellular origin. The cellular origin of viral *onc* sequences was first confirmed by molecular hybridization of cDNA probes made from Kirsten and Harvey sarcoma virus RNA to cellular nucleic acid¹⁵⁻¹⁷. However, biochemical definition of the relationship between the *onc* genes of these viruses and the cell was complicated by the fact that Harvey and Kirsten sarcoma viruses share sequences not only with their helper virus and the cell, but also with a defective endogenous retrovirus of rat cells¹⁵⁻¹⁸. The subsequent findings of cellular prototypes of the *src* gene of RSV¹⁹ and of the *onc* gene of Moloney sarcoma virus²⁰ provided much simpler systems for biochemical analyses because all specific sequences of these *onc* genes were present in normal cells. Homologues (proto-*onc*) of over a dozen other viral *onc* genes have since been found in normal cells³. Any vertebrate species may harbour prototypes of many known viral *onc* genes, including those of viruses that can only infect distantly related animals. For example, avian cells contain eight *onc* gene prototypes of two unrelated taxonomic groups of avian retroviruses (for example numbers 1-8 of the avian sarcoma-leukosis group^{3,6} and number 9 reticuloendotheliosis virus^{3,21} in Table 1), as well as partially related, but probably incomplete, prototypes of the *onc* genes of mammalian viruses such as murine Abelson leukaemia³⁰, Moloney sarcoma²⁰ and feline sarcoma virus²². Similarly, mammalian cells contain *onc* gene prototypes of mammalian retroviruses as well as genes related to *onc* genes that have only so far been found in avian retroviruses, for example those of avian MC29 virus^{23,24}, RSV²⁵, avian myeloblastosis virus²⁶ and Fujinami sarcoma virus (FSV)²⁷ but possibly not of reticuloendotheliosis virus²¹. Prototypes of the *onc* genes of RSV, MC29, Abelson and FSV have even been found in *Drosophila*^{27,28}. Thus, normal vertebrate cells contain many, but probably not all, known proto-*onc* genes.

Only with the advent of molecular cloning has the structural relationship between cellular proto-*onc* genes and viral *onc* genes become completely clear. Beginning with the cellular prototype of Moloney sarcoma virus²⁹, prototypes of most

known viral *onc* genes have since been cloned and compared with viral counterparts by heteroduplex analysis, endonuclease site mapping, fingerprinting and complete sequence analysis. Such comparisons have revealed that those parts of viral *onc* genes that are unrelated to essential genes of retroviruses have closely related counterparts in normal cells and, most importantly, that these counterparts in normal cells are not linked to any of the genes essential to retroviruses²⁹⁻³⁷.

Before considering in some detail the structural differences between viral *onc* genes and cellular proto-*onc* genes, it may be helpful to outline the kind of questions that preceded the analysis and the kind of answers that have emerged from it. The first question was why should the homology between the viral and cellular sequences be so close? The probable answer is that the close homology is a consequence of the short tenure of these sequences in retroviruses. As pointed out above, due to their nonessential, parasitic nature, these sequences are readily deleted and hence are allowed little evolutionary time for divergence from cellular progenitors.

The second question was why are cellular proto-*onc* genes transduced into retroviruses so infrequently? The answer appears to be that at least two illegitimate recombination events are necessary to transduce a cellular sequence into a retroviral vector, because cellular prototypes of *onc* genes and essential retroviral genes are unrelated. Probably, specific mutations are also necessary to convert a proto-*onc* gene precursor into an *onc* gene product. Moreover, these events would have to take place during the lifespan of one infected animal for a transmissible virus to emerge. The low probability of such an event is consistent with the rare occurrence of retroviruses with *onc* genes.

The third question was whether the data would favour the oncogene hypothesis or the provirus hypothesis. At first sight, the close sequence homologies between viral *onc* genes and cellular prototypes suggested that normal cells contain retroviral *onc* genes, in accord with the oncogene hypothesis. That, in turn, led to the quantitative or dosage model in which enhanced dosage of a normal gene is postulated to be sufficient to cause cancer. Consistent with this view, cellular prototypes of viral *onc* genes were named cellular (c) oncogenes³⁸ and referred to as enemies from within the cell³⁹. The alternative view^{1,6}, consistent with the provirus hypothesis, is the qualitative model in which vital differences exist between the structure and function of the viral *onc* genes and their cellular prototypes. These two models are not mutually exclusive, since many

Table 1 Identified *onc* genes and their viral prototypes

Oncogenes of retroviruses			Cancer in animal		Transformation in cell culture		
	No.	onc gene	Sarcoma	Carcinoma	Acute leukaemia	Fibroblasts	Blood cells
Avian viruses							
Rous sarcoma	1	src	+	—	—	+	?
Fujinami sarcoma	2	$\Delta gag-fps$	+	—	—	+	—
Yamaguchi sarcoma	3	$\Delta gag-yes$	+	—	—	+	—
Rochester-2 sarcoma	4	$\Delta gag-ros$	+	—	—	+	—
Myelocytomatosis MC29	5	$\Delta gag-myc$	+	+	+	+	+
Carcinoma MH2	6	$\Delta gag-mht/myc^*$	+	+	+	+	+
Avian erythroblastosis	7	$\Delta gag-erb^A/erb^B$	+	+	+	+	+
Myeloblastosis and erythroblastosis-E26	8	$\Delta gag-amv-\Delta env^\dagger$	—	—	+	—	+
Reticuloendotheliosis	9	$\Delta gag-amv-ets$ rel	—	—	+	—	+
Mammalian viruses							
Moloney murine sarcoma	10	mos	+	—	—	+	—
Harvey and Kirsten rat sarcoma	11	ras	+	—	+	+	?
Abelson murine leukaemia	12	$\Delta gag-abl$	—	—	+	+	+
FBJ murine osteosarcoma	13	fos	+‡	—	—	+	?
ST and GA feline sarcoma	14	$\Delta gag-fes$	+	+	—	+	—
SM feline sarcoma	15	$\Delta gag-fms$	+	—	—	+	—
Simian sarcoma	16	$\Delta env-sis$	+	—	—	+	—

Results are based on refs 1-4, 6, 18, 30 and others, see text. * N. C. Kan, C. S. Flordellis, C. F. Garon, P. D. and T. Papas (unpublished), see text. † *amv* is synonymous with the term *myb* in ref. 3. ‡ Osteosarcoma³.

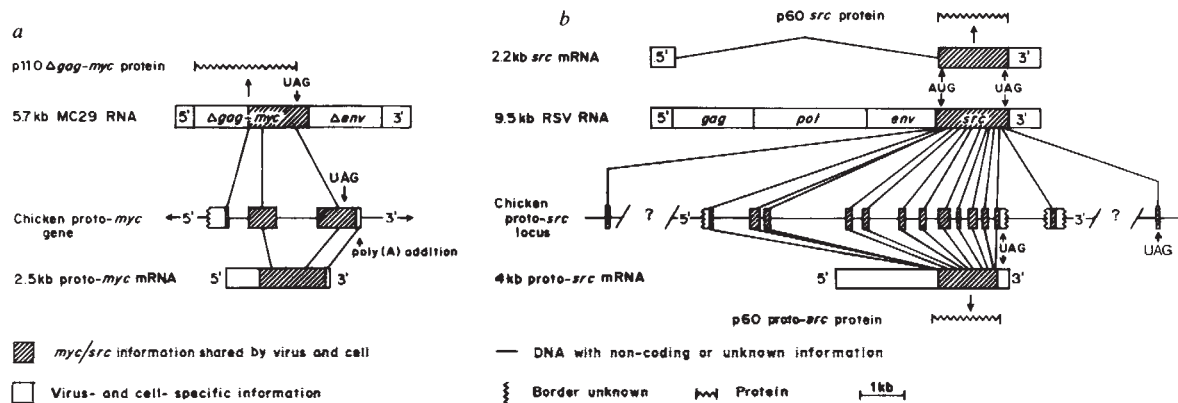


Fig. 1 Structural relationships between viral *onc* genes and cellular proto-*onc* genes. **a**, Juxtaposition of the hybrid (Δ gag-myc) *onc* gene of MC29, shown within the genetic structure of the virus, and the p110 Δ gag-myc protein with the proto-myc gene of the chicken and the 2.5-kb proto-myc mRNA. 5', 3' Represent terminal noncoding RNA sequences. *gag*, *pol* and *env* are complete and Δ partial complements of the three essential retroviral genes. The data are based on refs 14, 23, 32, 33 and 40. **b**, Juxtaposition of the *src* gene shown within the genetic map of Rous sarcoma virus (RSV), of the 2.2-kb *src* mRNA and the p60 *src* protein with the chicken proto-*src* locus, the known 4-kb proto-*src* mRNA and the putative p60 proto-*src* protein. The data are based on refs 8–11, 31, 42–47.

different mechanisms may lead to the same cancers. There is now ample sequence information to suggest differences between viral *onc* genes and proto-*onc* genes but still insufficient genetic and functional knowledge to determine whether cellular proto-*onc* genes are indeed cellular oncogenes; hence, the term proto-*onc* is used here rather than *c-onc* to emphasize this uncertainty. However, the following arguments suggest that most available data favour the qualitative model for the conversion of cellular sequences into oncogenic agents.

onc and proto-*onc* genes

Recent sequence comparisons reveal that many, perhaps all, viral *onc* genes differ from their cellular prototypes in some genetic information. This is particularly clear for the relationship between hybrid *onc* genes and their cellular prototypes. Such *onc* genes consist of a hybrid of two types of coding sequence, one derived from essential virion genes, typically a part of the *gag* gene (Δ gag), the other from a cellular locus (*x*). They have the generic Δ gag-*x* structure and are defined by a hybrid protein product⁶. The first hybrid *onc* gene was identified in avian MC29 virus⁴⁰. This gene has a 5' *gag* region of about 1.5 kilobases (kb) and a 3' region, termed *myc*, of 1.6 kb that is derived from the cell (Fig. 1). The Δ gag-myc gene, like the Δ gag-*x* genes of other retroviruses, such as Abelson leukaemia virus, feline sarcoma virus and avian FSV, all code for Δ gag-*x* proteins with probable transforming function^{3,6,30} (Table 1).

The *myc*-related gene of the chicken, termed proto-*myc*, was recently cloned, sequenced and compared with a DNA clone of the viral Δ gag-myc gene^{32,33} as shown schematically in Fig. 1a. Both DNAs share virtually identical *myc* coding regions with only nine nucleotide changes, corresponding to seven amino acid changes; there is an intron of possibly 0.45-kb and a 1-kb intron in the cellular *myc* sequence³³ (D. Watson, P.H.D. and T. Papas, unpublished). The genetic information of the 5' exon of the proto-*myc* gene continues in the 5' direction. The additional proto-*myc* information is approximately 0.9 kb based on size estimates of a 2.5-kb proto-*myc* mRNA²³. A small 10-nucleotide portion of this 5' exon corresponds to the 5' end of the *myc* sequence in MC29³³. The viral 110,000-molecular weight (110-kD) Δ gag-myc protein would share only the 3' 60-kD region with the yet unknown cellular proto-*myc* product. Assuming that specific structure has specific function, it is likely that the different 5' coding sequences of the viral and cellular *myc*-containing genes encode distinct functional domains.

The cellular homologue of the Δ gag-*fps* transforming gene of FSV also lacks Δ gag and differs from viral *fps* in point mutations³⁴. Moreover, preliminary sequence analysis of chicken proto-*fps* indicates that the coding sequence continues

past the 5' boundary of overlap with the viral *onc* gene (P. Seeburg, W-H. Lee and P.H.D., unpublished). The virus encodes a 140-kD gag-related protein¹³. A 98-kD protein presumably encoded by proto-*fps* has been observed in several types of normal cells⁴¹. The point mutations by which the viral *fps* differs from the cellular *fps* plus its Δ gag, possibly replacing cellular coding sequences of proto-*fps*, may be the reasons that the viral protein transforms and the cellular protein does not. The cellular 160-kD homologue of the 120-kD Δ gag-*x* protein of Abelson murine leukaemia virus is another example of major structural differences between a hybrid *onc* gene and its cellular prototype³⁰.

In the case of viral *onc* genes that are not hybrids with another viral gene but are made largely from cell-derived sequences, significant qualitative differences still emerge from a structural comparison of the *onc* genes and their cellular counterparts. For example, the *src* gene of RSV differs from its cellular prototype in several ways^{42,43} (see Fig. 1b). First, the coding regions of proto-*src* and *src* genes from different viral strains vary in scattered point mutations amounting to 1–2% of the sequence^{31,43}. Also, at both ends the *src* gene includes 0.1-kb regions that are not related to essential retroviral genes, appear not to be contiguous with the known proto-*src* locus, and are not part of the known 4-kb proto-*src* mRNA^{42–45}. The 5' boundary of overlap between the contiguous part of proto-*src* and viral *src* begins in a presumably noncoding exon of the proto-*src* locus that is distinct from the viral counterpart and is part of the 4-kb proto-*src* mRNA^{44,45}. Surprisingly, the intron between this and the next 3' proto-*src* exon is said to contain a transcriptional TATA promoter⁴³. From an AUG within the second exon, the coding region of *src* overlaps with proto-*src* for about 1.5 kb scattered in 11 exons over 7 kb of proto-*src*. In the 12th exon, *src* and proto-*src* cease to overlap. The reading frame and the mRNA of proto-*src* terminate with a unique coding sequence of 57 nucleotides that is not shared with viral *src*. By contrast, viral *src* terminates with a unique hybrid sequence of 39 nucleotides, 31 of which are derived from a region located 0.9 kb downstream from the 3' boundary of overlap between *src* and the major gene of the proto-*src* locus. The remaining eight nucleotides, including the UAG stop codon, are from an unknown source. Figure 1b shows that the termini of most genes of the proto-*src* locus are still unknown. Thus, proto-*src* and viral *src* are not isogenic, concordant with earlier work describing differences between *src* and proto-*src* protein products⁴⁶.

A model that could account for these differences and for the size differences between the 4-kb proto-*src* mRNA^{44,45} and the 2.2-kb viral *src* mRNA^{44,47} suggests that most of the coding sequence of viral *src* represents the 3' domain of the major proto-*src* gene. The 3' terminal coding region of viral *src* would

be derived from another gene of the proto-*src* locus located downstream and from another unknown cellular sequence. The discrepancy between the sizes of the mRNAs could then either reflect unique noncoding or coding exons of the major proto-*src* gene. If proto-*src* contains unique coding exons, the primary translational proto-*src* product would have to be reduced by processing, as the putative product of proto-*src* is approximately the same size as the 60,000 dalton viral *src* protein⁴⁶. Thus, the origin of viral *src* must have been the product of a complex process that involved transduction by a retrovirus, specific deletions, and multiple fusions of several cellular genetic elements via illegitimate recombinations.

The relationship between the transforming gene of simian sarcoma virus and its cellular prototype is another example of qualitative differences between a viral *onc* gene and its cellular progenitor. The reading frame of the cellular homologue of *sis*, the *onc*-specific sequence of simian sarcoma virus, continues in the 5' direction beyond the boundary of overlap with the viral gene⁴⁸. It also differs from viral *sis* in point mutations. Moreover, nucleic acid sequence analysis suggests that the coding sequence of the viral *sis* protein includes a specific region derived from essential virion genes (*Δenv*) which are not present in the cellular homologue⁴⁹. The coding sequence of the cellular homologue of *rel*, the *onc*-specific sequence of avian reticuloendotheliosis virus, also appears to extend over the boundaries of homology with the virus, and numerous point mutations set apart the viral and cellular *rel* sequences²¹. The viral *rel* sequence measures 1.4 kb whereas proto-*rel* mRNA measures 4 kb and contains cell-specific 3' sequences²¹. The *ras* gene of Harvey and Kirsten sarcoma viruses also differs from cellular proto-*ras* by point mutations⁵⁰⁻⁵². But there may be further differences between the 0.6-kb viral *ras* gene and proto-*ras*, since proto-*ras* encodes a 1.4- and a 5-kb mRNA species perhaps by alternate splicing⁵³. (Viral *ras* is thought to be translated from the 6-kb and 8-kb virion RNAs of Harvey and Kirsten virus⁵³.) Qualitative differences, including both point mutations and unique coding elements, have also been observed between the *onc*-specific sequence of Moloney sarcoma virus, termed *mos*, and its cellular equivalent with five unique 5' codons in viral *mos* and 21 in cellular *mos*^{35,54}. Since the *mos* sequence of the cellular gene terminates with a possible splice acceptor site at the 5' boundary of overlap with the viral *mos* sequence³⁵, it may contain a yet undefined 5' exon.

Similar situations exist for the relation of both FBJ-murine sarcoma virus⁵⁵ and avian myeloblastosis virus (AMV)^{36,56} to their cellular prototypes. In both cases the viral *onc* genes represent deletions and mutations of cellular genes. The specific sequence of AMV, termed *amv* or *myb*, measures 1.2 kb^{56,57} whereas the proto-*amv* gene is thought to have more genetic information³⁶; the viral *amv* protein is estimated at 48 kd whereas the proto-*amv* protein is 105 kd⁵⁸. A proto-*amv* mRNA of about 4 kb has been found in normal cells with unique cellular sequences 5' and 3' to the AMV-related region which could encode a 105-kd protein^{45,59,60}.

It will be particularly difficult to interpret possible functional relationships between a viral *onc* gene and a cellular prototype if an oncogenic virus is composed of elements from more than one cellular gene, as well as from essential retroviral elements. Recently we have found that avian erythroblastosis virus E26 contains an internal subset of the transformation-specific sequence of AMV in addition to another E26-transformation-specific sequence (*ets*) which are linked to *Δgag* to form a hybrid *Δgag-amv-ets* gene of three different genetic elements (Table 1) (ref. 61 and C. Moscovici and P. Duesberg, in preparation). Since the proto-*amv* gene terminates only 27 nucleotides downstream of the 3' border of overlap with *amv*³⁶ and since a proto-*amv* clone did not hybridize with *ets*, *ets* appears to be derived from a cellular gene that is different from proto-*amv*. That E26 predominantly causes erythroblastosis whereas AMV exclusively causes myeloblastosis may be related to which parts of proto-*amv* have found their way into each virus, to the additional alterations such as linkage to *Δgag*, and to *ets*. By

extrapolation, the much larger cellular proto-*amv* gene may have other unique functional domains.

Another case in point is the avian carcinoma virus MH2, which appears to contain possibly two different *onc* genes. A recent analysis of the genetic structure has shown that this virus contains a novel hybrid *onc* gene in which *Δgag* is linked to a unique 1.3-kb sequence termed *mht* (MH2-transformation-specific) (N. C. Kan, C. S. Flordellis, C. F. Garon, P. Duesberg and T. S. Papas, unpublished, and H. W. Jansen, T. Patschinsky and K. Bister, personal communication). This gene appears to encode the *gag*-related p100 protein of MH2, which is probably necessary for transformation. Based on hybridization with chicken DNA and cloned proto-*myc* DNA, *mht* appears to derive from a cellular gene that is different from proto-*myc*. MH2 contains, in addition, the *myc* sequence which it shares with MC29, CMII and OK10. Considering the unique genetic structure of MH2 (5' *Δgag*(2 kb)-*mht*(1.3 kb)-*myc*(1.3 kb)-*Δenv*(?)-(0.7 kb)3') and the size of the *Δgag-mht* gene as defined by the *gag*-related p100 protein, it appears that the *myc* sequence of MH2 is part of a second gene that is different from the *Δgag-myc* genes of MC29, CMII and OK10. Thus, oncogenic function of MH2 may be the product of two genes, one composed from viral *gag* and proto-*mht*, the other from proto-*myc* and perhaps also from essential retroviral elements present in MH2. Further work will be necessary to determine whether both of these genes have to cooperate for viral oncogenicity or whether each, by itself, has oncogenic function.

The *gag*-linked, transformation-specific sequence of avian erythroblastosis virus (AEV) is significantly larger (3 kb) than the *gag*-related 75-kd protein product encoded by the virus^{6,11,9}. Therefore, AEV also has the potential to encode an additional transformation-specific product. Indeed, it has been proposed that AEV contains two genes, *Δgag-erb^A* and *erb^B* (Table 1) which are composed of genetic elements from the retrovirus *gag* and from two different cellular genes¹²⁰. However, a recent genetic analysis has concluded that *erb^B* is sufficient for oncogenicity and that *Δgag-erb^A* may only serve as an ancillary function¹²⁰.

From these data, it would appear that viral *onc* genes and cellular proto-*onc* genes are not isogenic. Available structural evidence favours the hypothesis that either deletion and mutation of cellular sequences or the addition of virus-specific information, such as *Δgag*, or both are needed to confer transforming function on proto-*onc* sequences transduced by viruses. In some cases, elements from several cellular genes appear necessary to generate an oncogenic virus. That is to say, qualitative changes are the norm.

Function of proto-*onc* genes

The normal cellular function of proto-*onc* genes is still unclear. The high degree of conservation of some proto-*onc* genes among animals suggests that these serve basic functions. The quantitative model, which postulates functional equivalence between proto-*onc* genes and viral *onc* genes, predicts that proto-*onc* genes are latent cancer genes. These may be activated by transduction into retroviruses or by acquiring promoters from adjacently integrated retroviruses or from other cellular genes by translocation or by amplification of the gene with its own promoter. The qualitative model would predict that mutations or rearrangements are necessary to render proto-*onc* genes oncogenic. To test these predictions, experimental transduction has been explored and transforming function of modified or unmodified proto-*onc* DNA has been tested by an assay that measures morphological transformation of the preneoplastic mouse NIH 3T3 cell line. Indirect assays have sought to correlate expression or alteration of proto-*onc* genes with neoplastic transformation.

Does transduction activate or generate transforming genes? Once the oncogene hypothesis was based on structural evidence, proto-*onc* genes were seen as functional and transducible equivalents of viral *onc* genes. Yet precursor/product analyses showed that in all of these cases, proto-*onc* genes had

been altered once they had become part of oncogenic viruses. If these structural alterations are functionally relevant, sequence transduction would only be the first step in the generation of a viral *onc* gene. The complete process must be accompanied by specific deletions and mutations. Hence, the low frequency of natural transduction probably reflects not only the lack of primary structural homology between transducing retroviruses and proto-*onc* genes but also the lack of functional homology which necessitates specific secondary alterations of proto-*onc* genes. Thus, natural transduction does not prove functional homology between *onc* genes and proto-*onc* genes. The extremely low frequency of these transductions precludes a reductionistic analysis. Analysis is limited to historical reconstruction of singular events that generated the original oncogenic viruses in animals infected by retroviruses (Table 1)^{2,4}.

The only experimentally reproducible transduction system reported to date is the apparent ability of partial (>75%) *src* deletion mutants to regain transforming function within about 2 months in 50% of infected chickens⁶². Regarding the relevance of this and other systems to natural transduction of proto-*onc* genes, it is important to distinguish between two different types of *onc* gene deletion mutants; those in which internal *onc* regions are deleted and those in which *onc* termini are deleted. Recovery of *onc* sequences by internal *onc* deletion mutants would not be a model for natural transduction, since natural retroviruses without *onc* genes lack *onc* termini. Such deletions may be repaired by homologous asymmetrical recombination with proto-*onc* genes, a process known to occur in retroviruses⁶³. Transduction of proto-*src* sequences by internal *src* deletion mutants of RSV has been reported to occur in an experimentally reproducible fashion^{62,64} and, in a single case, transduction of proto-*myc* sequences by an internal *myc*-deletion of MC29 has been reported⁶⁵. Since these deletion mutants were not molecularly cloned, it has not been excluded that cross-reactivation with unappreciated contaminants may have contributed to the recovery of these *onc* sequences.

By contrast, experimental transduction of proto-*onc* genes by deletion mutants lacking a complete *onc* gene terminus would be relevant to the natural origin of retroviruses with *onc* genes because they would have to go through at least one illegitimate recombination with proto-*onc* in order to regain a complete *onc* gene³¹. It has been reported that RSV deletion mutants that lack an entire 5' (refs 62, 64) or 3' *src* terminus⁶⁶ regain *src* function from proto-*src* at essentially the same frequency as internal *src* deletions⁶². Nevertheless, several observations cast doubt on the transductional origin of regained *src* termini in this system: (1) It remains unexplained why these terminal *src* deletions were reported to regain *src* at essentially the same frequency as internal *src* deletions^{62,64}, as recovery of *src* termini by illegitimate recombination would be a much more complex and hence a less frequent process than recovery of an internal *src* sequence by homologous recombination. (2) It has not been excluded that essential regions, in particular both terminal *src* regions, may have pre-existed in the transducing virus, as the *src* deletions used were not molecularly cloned and the regained *src* genes differed from parental equivalents only in point mutations but not in proto-*src*-specific sequences^{31,43}. Recent evidence that the terminal sequences of *src* differ in different RSV strains^{42,43} suggests an assay to determine whether a complete *src* terminus was transduced from proto-*src* or pre-existed in the virus stock. *src* termini regained *de novo* from proto-*src* would be expected to show the same variations as those in different RSV strains.

Taken together, available experimental systems do not illuminate the mechanism of proto-*onc* sequence transduction by natural retroviruses that leads to viral *onc* genes.

In vitro transformation by proto-*onc* DNA ligated to retroviral promoters: DNA of the cellular prototypes of the Moloney and Harvey (Ha) sarcoma viruses, termed proto-*mos* and proto-*ras* (Ha), were shown to transform mouse 3T3 cells if ligated with promoter sequences of Moloney and Harvey sarcoma viruses⁶⁷⁻⁶⁸. In subsequent experiments, the long terminal

repeat was directly shown to function as a promoter⁶⁹. These experiments appear to favour the quantitative model—albeit with reservations. Specific transforming activities of the sarcoma virus promoter-proto-*mos*/*-ras* constructions are 10–200-fold lower in 3T3 cells than those of DNA clones of authentic viral *onc* genes^{30,67,68,70}. This may reflect yet unappreciated differences between viral *onc* genes and cellular prototypes (such as a unique cellular *mos* exon, see above). Another indication of such differences is that the viral *mos* DNA, but not proto-*mos* DNA, transforms 3T3 cells even with a viral promoter at its 3' end^{30,67}. That promoter-activated proto-*mos* and proto-*ras* have never been found in natural cancers although leukaemia viruses which could provide promoters are quite ubiquitous in mice makes one wonder whether such structures are sufficient to transform cells in the animal. The view that activated proto-*ras* may not be sufficient to transform cells in the animal is in good agreement with the observations that expression of proto-*ras* is elevated in several normal conditions, for example, in developing mouse embryo cells⁷¹ and in regenerating rat liver cells⁷². The ultimate test of functional homology between viral *onc* genes and cellular prototypes will be possible when a Moloney or Harvey virus can be reconstructed from the appropriate proto-*onc* gene and retroviral vector sequences and tested for its ability to induce sarcomas in mice.

Except for proto-*ras* and proto-*mos*, all other proto-*onc* sequences from normal cells ligated to retroviral promoters have been consistently negative in the 3T3 cell transformation assay. Those tested have included proto-*myc*, proto-*mos* from humans⁷³ and proto-*src*⁷⁴ which failed to transform even though the proto-*src* was expressed in 3T3 cells (T. Robins and G. Vande Woude; G. Cooper; R. Parker and J. M. Bishop, personal communications).

In vitro transformation by mutated proto-*onc* genes: Analysis of DNA from human bladder carcinoma cell lines has identified point mutated proto-*ras* that transforms 3T3 cells. The single base change affects the 12th codon of the p21 *ras* gene product⁷⁵⁻⁷⁸. The mutation does not cause overproduction of *ras* in the cell line⁷⁵. Although consistent with the qualitative model, this does not prove that mutated proto-*ras* is carcinogenic in animals. As yet this mutation has not been found in over 50 primary human carcinomas, including 10 bladder, 9 colon and 10 lung carcinomas⁷⁹ and 14 additional bladder and 9 kidney carcinomas (R. Muschel and G. Khoury, personal communication). These results have an obvious bearing on the question of whether the 3T3 cell assay detects primary cancer genes or secondary mutations that evolve during tumour progression or cultivation of tumour cells in tissue culture^{78,80,81}. An example of increasing transforming potential of proto-*ras* upon subsequent transplantation of the primary carcinomas has recently been described¹²¹. This result is consistent with the view that the transforming function of *ras* DNA may be a consequence rather than the primary cause of the tumour. A qualitatively altered proto-*mos* sequence that transforms 3T3 cells has also been detected in a mouse myeloma line⁸². About one-quarter of the proto-*mos* coding sequence had been replaced by a transposon-like sequence presumably as a result of transposition and deletion. Again, this does not prove that the altered proto-*mos* caused the myeloma because altered proto-*mos* was associated with only 2 out of 20 myeloma cell lines and primary tumours have not been tested (E. Canaani, personal communication).

Proto-*onc* expression and carcinogenesis: Enhanced expression of the cellular proto-*myc* has been postulated to be the cause of chicken B-cell lymphoma^{83,84}. The lymphoma is a clonal cancer^{83,84} that is caused in a small fraction of chickens infected by avian leukosis virus (which has no *onc* genes) after latent periods of over 6 months²⁻⁴. Proto-*myc* (see Fig. 1) is thought to be activated by the promoter of a retrovirus integrated nearby (within 5 kb)^{83,84} analogous to the model of downstream promoter activation by bacterial transposons⁸⁵. Consistent with the hypothesis, elevated expression of viral proto-*myc* hybrid mRNAs, ranging in size from 2.5 to over 5 kb, was observed

in 80% of the retroviral chicken lymphomas investigated^{83,84}. The hypothesis is that activation of proto-*myc* in this way is equivalent to the effects of viral Δ gag-*myc* from MC29 integration. However, the hypothesis is inconsistent with the disease for the following reasons: (1) Promoter insertion and elevated expression of proto-*myc* DNA are not found in all chicken B-cell lymphomas^{83,84} and, therefore, it is not proven that activation of proto-*myc* is necessary for this cancer. (2) The phenotype of the disease is different from that of each of the multiple cancers caused by MC29^{3,6}. (It may be argued that identical biochemical mechanisms of proto-*myc* and MC29 produce different cancers in different target cells but this fails to explain why retroviruses that cause lymphoma do not cause MC29-specific cancers at the same rate.) (3) The hypothesis is inconsistent with the latent period and the clonality of the cancer by the following calculations. The titre of retroviruses in chicken is high ($>10^7$ per ml serum), the number of target cells is also very high ($>10^7$)^{2,4}, and integration by retroviruses is not site specific³. Since the complexity of the chicken genome is about 10^6 kb, the critical carcinogenic integration within 5 kb of proto-*myc* should occur in one out of 2×10^5 infected lymphoblasts. Thus, if all lymphoblasts are susceptible to transformation, the latent period in the animal should be short, similar to that of MC29, and the tumours should not be clonal. If it is argued that integration near proto-*myc* is a particularly rare event but necessary and sufficient for lymphomagenesis, some animals should get the disease immediately after infection and others after longer latent periods. However, short latent periods are not observed in avian viral leukaemia^{2,4}. It would appear that an unknown critical event determines clonality and perhaps the long latent period of the disease. (4) When transforming function of the DNA from B-cell lymphomas was assayed in 3T3 cells, *myc*-related DNA was not detected, even though its presumed functional equivalent, the Δ gag-*myc* gene of MC29, is capable of transforming 3T3^{78,86} and other rodent cells⁸⁷. Instead, another DNA sequence was identified by the assay^{88,89}.

Based on these 3T3 cell transformation results it has been suggested that proto-*myc* has a transient initiation function that generates a lymphoma maintenance gene during the long latent period. This gene may be the DNA that transforms 3T3 cells. This puts the putative function of proto-*myc* at odds with the function of the Δ gag-*myc* gene of MC29. It suggests that activation of proto-*myc* may not be the critical carcinogenic event and that MC29 may not be an adequate model for the function of proto-*myc*. That would explain why the MC29 *onc* gene initiates and directly maintains the carcinomas, sarcomas and acute leukaemias caused within 1–2 weeks of infection with virus⁶ whereas lymphomas follow proto-*myc* activation only after long latency. Our evidence that the chicken proto-*myc* gene and the *onc* gene of MC29 have different primary structures and genetic information suggests that each gene may have unique functional properties. The activated normal or altered proto-*myc* genes may then have a role in lymphomagenesis by a different mechanism from that of MC29.

To assess whether proto-*myc* has a specific or aetiological role in this tumour, it would be necessary to understand the function of normal proto-*myc*. It would also be important to determine whether other genes are simultaneously activated. It is also conceivable that proto-*myc* activation is a consequence rather than a cause of the disease. Consider that an unknown primary carcinogenic event activates proto-*myc* in prospective lymphoma cells because proto-*myc* encodes a fundamental function, perhaps for cell division. By integrating into the active proto-*myc* during this stage, the retrovirus may then maintain elevated proto-*myc* expression (and perhaps alter the gene) in a majority but not in all lymphoma cells.

It is interesting that similar experiments designed to detect activation of cellular proto-*onc* genes in murine^{90,91} and bovine⁹² viral leukaemias have not confirmed the activation hypothesis of known cellular proto-*onc* genes. Nevertheless, integration sites for murine leukaemia virus in rat⁹⁰ and mouse⁹¹ DNA were not completely random. Of 16 tumours, 5 had

proviruses in a common region⁹⁰.

Most other efforts to relate elevated expression of proto-*onc* genes with neoplastic transformation have failed to show reproducible positive correlations. Expression of proto-*onc* genes was observed in several forms of neoplastic transformation^{26,93}. However, expression of the cellular prototypes of *src*^{44,45}, *ras*^{71,72,94}, *myc*^{23,45,71,95}, *fps*⁴¹, *rel*²¹, *abl*^{30,71} and of avian erythroblastosis⁴⁵, myeloblastosis^{59,60} and FBJ murine sarcoma virus⁷¹ were observed in normal cells. This implies that expression of proto-*onc* genes is not sufficient for neoplastic transformation.

Proto-*onc* translocation and carcinogenesis: Translocation has been proposed as a mechanism to activate cellular cancer genes^{96,97}. The chromosomal locus that carries human and mouse proto-*myc* is translocated to one of several chromosomes that carry immunoglobulin genes in most mouse plasmacytomas and in human B-cell lymphoma lines, derived from Burkitt or non-Burkitt-type lymphomas^{98–105}. In man, proto-*myc* is located on chromosome 8 and is typically translocated to immunoglobulin loci of chromosome 14 and less frequently of chromosomes 2 or 22^{99,106}. Translocation is suspected to activate proto-*myc* to function as a cellular oncogene in these tumours^{98,101–106}. The cross-over point with the recipient chromosomes varies in different tumours, but the translocation appears to recombine, in most cases, a region 5' to proto-*myc* with a region 5' to an immunoglobulin gene^{98,99,101,103}. In many but not all human and mouse lymphomas, the cross-over is within a few kilobases of the proto-*myc* sequence. In these cases, the proto-*myc* locus becomes rearranged, which is reflected by length polymorphism of *myc*-related DNA fragments generated by restriction endonucleases^{98–103}. It seems unlikely, however, that new hybrid genes would be formed, since the immunoglobulin and *myc* coding regions are joined in opposite transcriptional orientation. Consistent with such rearrangement, proto-*myc* mRNAs in some tumour lines are smaller than normal counterparts^{98,101}. There is no consensus at this time as to whether proto-*myc* expression is enhanced as a consequence of translocation. Some investigators report elevated expression compared with other B-lymphoblasts or lines^{98,104}, while others report essentially normal levels of mRNA (refs 93, 101 and G. Cooper, personal communication). As with the chicken B-cell lymphoma, the proto-*myc* DNA from mouse lymphomas was not found to transform 3T3 cells^{98,101,105}. However, a 3T3 cell transforming DNA related to the DNA that transforms 3T3 cells in chicken B-cell lymphomas⁸⁹, was recently identified in Burkitt lymphoma cells (G. Cooper, personal communication).

Clearly, the correlations with proto-*myc* translocations must be relevant to these lymphomas in some way. However, an aetiological role of proto-*myc* in these lymphomas is even more difficult to prove than in the avian viral lymphomas, since there is neither a mammalian retrovirus with a *myc*-related *onc* gene for comparison nor a consistent correlation with increased proto-*myc* expression. To determine whether proto-*myc* translocation is the cause or a consequence of the cancer, it would be necessary to prove that this translocation is indeed the primary malignant event in these mammalian lymphomas. It would then also become necessary to exclude that the reciprocal translocation of the displaced immunoglobulin locus into the locus vacated by proto-*myc* may be relevant to lymphomagenesis.

Moreover, translocations associated with human lymphomas do not always involve the proto-*myc* locus. In some non-Burkitt lymphomas, the immunoglobulin locus of chromosome 14 recombines with unknown loci from other chromosomal donors, in particular chromosome 18¹⁰⁷. This suggests that recombination of an immunoglobulin locus with one of several loci, that must not be proto-*myc*, is an essential event in lymphomagenesis and that translocation of a proto-*onc* gene may not be necessary for lymphomagenesis.

Another translocation is that which accounts for the variant chromosome 22, the Philadelphia chromosome, in the majority

of chronic human myeloid leukaemias. This results from a reciprocal translocation between chromosomes 22 and (frequently) 9¹⁰⁶. Chromosome 22 carries the λ light-chain immunoglobulin locus. In several such cases, proto-*abl*, the human cellular prototype of the Abelson murine leukaemia virus *onc* gene, is translocated from chromosome 9 to 22¹⁰⁸. This translocation is linked to enhanced expression and amplification of proto-*abl* sequences in the myeloid leukaemia cell line K-562 (S. Collins and M. Groudine, personal communication). However, with regard to the question of whether proto-*abl* translocation may be necessary for the leukaemia, it has been pointed out that variant Philadelphia chromosomes with translocated elements from 18 different chromosomes have been observed⁹⁹. The view that proto-*abl* translocation may be the cause of the leukaemia is also challenged by the observation that the translocation may be a late event which also occurs in 'normal' lymphocytes derived from the same clone as the leukaemic cells¹⁰⁹.

Proto-*onc* amplification and cancer: Amplification and elevated expression of proto-*myc* sequences have been observed in a human promyelocytic leukaemia^{110,111} and in a colon cancer cell line¹¹². Since amplified sequences are a specific phenomenon associated with the viability of certain cell lines in culture¹¹³, it has been suspected that proto-*onc* gene amplification may be a critical event in carcinogenesis^{110,112}. In the promyelocytic leukaemia cells, *myc* gene amplification was observed in stored, uncultured cells as well as in a derivative cell line, HL60¹¹⁰, but it is not known whether the *myc* amplification was restricted only to leukaemic cells of this patient. In HL60 cells, amplified proto-*myc* sequences have been localized to double minute chromosome structures¹¹¹. A 3T3 cell transforming DNA has been isolated from HL60 cells¹¹⁴ that was recently shown to be unrelated to proto-*myc* but distantly related to the *ras* genes of Harvey and Kirsten virus (S. Collins, personal communication). However, *myc* amplification has not been observed in 10 fresh myeloid leukaemia cases examined (S. Collins, personal communication) and in other myeloid leukaemia cell lines including K-562¹¹¹. Instead, proto-*abl* seems to be amplified in K-562 cells (S. Collins, personal communication).

In addition, amplification of proto-*ras* has been observed in a mouse adrenocortical tumour line¹¹⁵. However, 6–10-fold amplifications of the proto-*ras* (Ha) gene, that do not coincide with neoplastic transformation, have been observed in certain strains of normal mice and hamsters¹¹⁶. It appears, therefore, that there is no consistent correlation between proto-*onc* gene amplification and carcinogenesis.

Proto-*onc* genes as potential cancer genes

Clearly, the identification of *onc* genes has moved viral carcinogenesis from the romantic into an academic age. The same cannot be said for the suspected role of proto-*onc* genes in cancer. As yet there is no functional and no consistent circumstantial evidence that proto-*onc* genes directly initiate and maintain cancer like the *onc* genes of retroviruses. There is also no such evidence that proto-*onc* genes encode one of the multiple initiation and promotion events that create virus-negative cancers^{81,96,117,118}.

The existential question for tumour virologists now is whether proto-*onc* genes are more relevant to cancer than other cellular genes. A statistical argument suggests that this may be the case. Since illegitimate recombination seems to be the *modus operandi* by which retroviruses transduce sequences from cells to make *onc* genes, any cellular gene could serve as a potential candidate for transduction. However, to date, not more than 16 different retroviral *onc* genes have been found. Moreover, some of these have apparently been isolated repeatedly, for example, the four isolates of the MC29 subgroup of avian acute leukaemia viruses, and the members of the avian Fujinami sarcoma virus subgroup⁶, and the Harvey and Kirsten sarcoma viruses⁵⁰ each share closely related *onc* genes but have different genetic structures, consistent with independent transductions.

Therefore, it is likely that cellular prototypes of retroviral *onc* genes are indeed more relevant to cancer than other cellular sequences—at least in the presence of retroviruses which could serve as appropriate mutagenic vectors. Although retroviruses which lack *onc* genes are ubiquitous in many animal species, it is rarely that retroviruses with *onc* genes have been isolated^{2–4}.

It is conceivable, however, that retroviruses without *onc* genes would generate cellular/viral hybrid *onc* genes without generating oncogenic viruses. One illegitimate recombination, for example, between retroviral *gag* and proto-*myc*, would be sufficient to form an MC29-type cellular *onc* gene, which could not leave the cell as an infectious virus. Cancers caused by such hybrid transforming genes may be much more frequent than the rare tumours caused by retroviruses with *onc* genes, since only one illegitimate recombination is necessary compared with the two that are necessary to generate oncogenic viruses.

Further work will be necessary to determine whether proto-*onc* genes are more likely than other genes to become carcinogenic in the absence of retroviruses. The known structural differences between viral *onc* genes and cellular proto-*onc* genes provide clues to how this could happen. Qualitative changes, such as deletions, point mutations and recombination with other cellular genetic elements, including transposable elements or translocated genes, are alterations that may confer oncogenicity to proto-*onc* genes. Such alterations might directly affect the function or might alter substrate specificity or the topography of the proto-*onc* gene products in the cell. Nevertheless, so far retroviral *onc* genes are the only proof that altered proto-*onc* genes can be cancer genes. There is as yet no conclusive example that an unaltered proto-*onc* gene can function as a cancer gene simply through enhanced transcription or gene amplification. Complete genetic definition and assays for the biological function of normal and mutated proto-*onc* genes are now necessary to understand their possible role in carcinogenesis.

I thank S. Aaronson, M. Baluda, J. M. Bishop, M. Botchan, M. Carey, S. Collins, G. Cooper, A. Feinberg, R. Gallo, H. Hanafusa, W. Kirsten, M. Krieger, W.-H. Lee, G. S. Martin, M. Nunn, T. Papas, T. Robins, H. Rubin, I. Verma, L.-H. Wang, D. Weaver and W. Phares for their helpful comments and discussions and for encouragement or for communicating results before publication. Preparation of the manuscript by Linda Brownstein and of the figure by B. Ricco are gratefully acknowledged. Research is supported by NIH grant CA 11426 from the NCI.

1. Duesberg, P. H. *Cold Spring Harb. Symp. quant. Biol.* **44**, 13–29 (1980).
2. Tooze, J. *The Molecular Biology of Tumor Viruses* (Cold Spring Harbor Laboratory, New York, 1973).
3. Weiss, R. A., Teich, N. M., Varmus, H. & Coffin, J. M. *The Molecular Biology of Tumor Viruses* (Cold Spring Harbor Laboratory, New York, 1982).
4. Gross, L. *Oncogenic Viruses* (Pergamon, New York, 1970).
5. Huebner, R. J. & Todaro, G. J. *Proc. natn. Acad. Sci. U.S.A.* **64**, 1087–1094 (1969).
6. Bister, K. & Duesberg, P. H. in *Advances in Viral Oncology* Vol. 1 (ed. Klein, G.) 3–42 (Raven, New York, 1982).
7. Temin, H. M. *J. natn. Cancer Inst.* **46**, 3–7 (1971).
8. Duesberg, P. H. & Vogt, P. K. *Proc. natn. Acad. Sci. U.S.A.* **67**, 1673–1680 (1970).
9. Martin, G. S. & Duesberg, P. H. *Virology* **47**, 494–497 (1972).
10. Duesberg, P. H. & Vogt, P. K. *Virology* **54**, 207–219 (1973).
11. Lai, M. M. C., Duesberg, P. H., Horst, J. & Vogt, P. K. *Proc. natn. Acad. Sci. U.S.A.* **70**, 2266–2270 (1973).
12. Maisel, J., Klement, V., Lai, M. M. C., Ostertag, W. & Duesberg, P. H. *Proc. natn. Acad. Sci. U.S.A.* **70**, 3536–3540 (1973).
13. Lee, W.-H. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 2018–2022 (1980).
14. Duesberg, P. H., Bister, K. & Vogt, P. K. *Proc. natn. Acad. Sci. U.S.A.* **74**, 4320–4324 (1977).
15. Scolnick, E. M., Rands, E., Williams, D. & Parks, W. P. *J. Virol.* **12**, 458–463 (1973).
16. Scolnick, E. M. & Parks, W. P. *J. Virol.* **13**, 1211–1219 (1974).
17. Tsuchida, N., Gilden, R. V. & Hatanaka, M. *Proc. natn. Acad. Sci. U.S.A.* **71**, 4503–4507 (1974).
18. Scolnick, E. M. *Microbiol. Rev.* **45**, 1–8 (1981).
19. Stehelin, D., Varmus, H. E., Bishop, J. M. & Vogt, P. K. *Nature* **260**, 170–173 (1976).
20. Frankel, A. & Fischinger, P. J. *Proc. natn. Acad. Sci. U.S.A.* **73**, 3705–3709 (1976).
21. Chen, I. S. Y., Wilhelmson, K. C. & Temin, H. M. *J. Virol.* **45**, 104–113 (1983).
22. Frankel, A. E., Gilbert, H. J., Porzig, K. J., Scolnick, E. M. & Aaronson, S. A. *J. Virol.* **30**, 821–827 (1979).
23. Vennstrom, B., Sheiness, D., Zabielski, J. & Bishop, J. M. *J. Virol.* **42**, 773–779 (1982).
24. Colby, W. W., Chen, E. Y., Smith, D. H. & Levinson, H. D. *Nature* **30**, 722–725 (1983).
25. Spector, D. H. *et al. Cell* **13**, 371–379 (1978).
26. Westin, E. H. *et al. Proc. natn. Acad. Sci. U.S.A.* **79**, 2194–2198 (1982).
27. Lee, W.-H., Liu, C.-P. & Duesberg, P. J. *Virol.* **44**, 401–412 (1982).
28. Shilo, B.-Z. & Weinberg, R. A. *Proc. natn. Acad. Sci. U.S.A.* **78**, 6789–6792 (1981).
29. Oskarsson, M. K., McClements, W. L., Blair, D. S., Maizel, J. V. & Vande Woude, G. F. *Science* **207**, 1222–1224 (1980).

30. Goff, S. P. & Baltimore, D. in *Advances in Viral Oncology* Vol. 1 (ed. Klein, G.) 127-139 (Raven, New York, 1982).
31. Lee, W.-H., Nunn, M. & Duesberg, P. H. *J. Virol.* **39**, 758-776 (1981).
32. Robins, T., Bister, K., Garon, C., Papas, T. & Duesberg, P. *J. Virol.* **41**, 635-642 (1982).
33. Watson, D. K., Reddy, E. P., Duesberg, P. H. & Papas, T. S. *Proc. natn. Acad. Sci. U.S.A.* **80**, 2146-2150 (1983).
34. Lee, W.-H., Phares, W. & Duesberg, P. H. *Virology* (in the press).
35. Van Beveren, C. *et al. Nature* **289**, 258-262 (1981).
36. Klempnauer, K.-H., Gonda, T. J. & Bishop, J. M. *Cell* **31**, 453-463 (1982).
37. Shalloway, D., Zelenetz, A. D. & Cooper, G. M. *Cell* **24**, 531-542 (1981).
38. Coffin, J. M. *et al. J. Virol.* **40**, 953-957 (1981).
39. Bishop, J. M. *Cell* **23**, 5-6 (1981).
40. Mellon, P., Pawson, A., Bister, K., Martin, G. S. & Duesberg, P. H. *Proc. natn. Acad. Sci. U.S.A.* **75**, 5874-5878 (1978).
41. Mathey-Prevot, B., Hanafusa, H. & Kawai, S. *Cell* **28**, 897-906 (1982).
42. Schwartz, D. E., Tizard, R. & Gilbert, W. *Cell* **32**, 853-869 (1983).
43. Takeya, T. & Hanafusa, H. *Cell* **32**, 881-890 (1983).
44. Wang, S. Y., Hayward, W. S. & Hanafusa, H. *J. Virol.* **24**, 64-73 (1977).
45. Gonda, T. J., Sheiness, D. K. & Bishop, J. M. *Molec. cell. Biol.* **2**, 617-624 (1982).
46. Smart, J. E. *et al. Proc. natn. Acad. Sci. U.S.A.* **78**, 6013-6017 (1981).
47. Mellon, P. & Duesberg, P. *Nature* **270**, 631-634 (1977).
48. Josephs, S. F., Dalla Favera, R., Gelmann, E. P., Gallo, R. C. & Wong-Staal, F. *Science* **219**, 503-505 (1983).
49. Devare, S. G., Reddy, E. P., Law, J. D., Robbins, K. C. & Aaronson, S. A. *Proc. natn. Acad. Sci. U.S.A.* **80**, 731-735 (1983).
50. Ellis, R. W., Lowy, D. R. & Scolnick, E. M. in *Advances in Viral Oncology* Vol. 1 (ed. Klein, G.) 107-126 (Raven, New York, 1982).
51. Dhar, R. *et al. Science* **217**, 934-936 (1982).
52. Capon, D. J., Chen, E. Y., Levinson, A. D., Seeburg, P. H. & Goeddel, D. V. *Nature* **302**, 33-37 (1983).
53. Ellis, R. W., DeFeo, D., Furth, M. E. & Scolnick, E. M. *Molec. cell. Biol.* **2**, 1339-1345 (1982).
54. Reddy, E. P. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 5234-5238 (1980).
55. Van Beveren, C., Van Straaten, F., Curran, T., Müller, R. & Verma, I. M. *Cell* **32**, 1241-1255 (1983).
56. Rushlow, K. E. *et al. Science* **216**, 1421-1423 (1982).
57. Duesberg, P. H., Bister, K. & Moscovici, C. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5120-5124 (1980).
58. Boyle, W. J., Lipsich, J. S., Reddy, E. P. & Baluda, M. A. *Proc. natn. Acad. Sci. U.S.A.* **80**, 2834-2838 (1983).
59. Chen, J. H. *J. Virol.* **36**, 162-170 (1980).
60. Gonda, T. J. & Bishop, J. M. *J. Virol.* **46**, 212-220 (1983).
61. Bister, K. *et al. Proc. natn. Acad. Sci. U.S.A.* **78**, 3677-3681 (1982).
62. Hanafusa, H., Halpern, C. C., Buchhagen, D. C. & Kawai, S. *J. exp. Med.* **145**, 1735-1737 (1977).
63. Martin, G. S., Lee, W.-H. & Duesberg, P. H. *J. Virol.* **36**, 591-594 (1980).
64. Wang, L.-H., Halpern, C. C., Nadel, M. & Hanafusa, H. *Proc. natn. Acad. Sci. U.S.A.* **75**, 5812-5816 (1978).
65. Ramsay, G. M., Enrietto, P. J., Graf, T. & Hayman, M. J. *Proc. natn. Acad. Sci. U.S.A.* **79**, 6885-6889 (1982).
66. Vigne, R., Breitman, M., Moscovici, C. & Vogt, P. K. *Virology* **93**, 413-426 (1979).
67. Blair, D. G. *et al. Science* **212**, 941-943 (1981).
68. Chang, E. H., Furth, M. E., Scolnick, E. M. & Lowy, D. R. *Nature* **297**, 479-483 (1982).
69. Kriegler, M. & Botchan, M. *Molec. cell. Biol.* **3**, 325-339 (1983).
70. DeFeo, D. *et al. Proc. natn. Acad. Sci. U.S.A.* **78**, 3328-3332 (1981).
71. Müller, R., Harmon, D. J., Tremblay, J. M., Cline, M. J. & Verma, I. M. *Nature* **299**, 640-644 (1982).
72. Gorgette, M., Petropoulos, C. J., Shank, P. R. & Fausto, N. *Science* **219**, 510-512 (1983).
73. Watson, R., Oskarsson, M. & Vande Woude, G. F. *Proc. natn. Acad. Sci. U.S.A.* **79**, 4078-4082 (1982).
74. Takeya, T. & Hanafusa, H. *J. Virol.* **44**, 12-18 (1982).
75. Tabin, C. J. *et al. Nature* **300**, 143-149 (1982).
76. Reddy, E. P., Reynolds, R. K., Santos, E. & Barbacid, M. *Nature* **300**, 149-152 (1982).
77. Der, J. C., Krontiris, T. G. & Cooper, G. M. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3637-3640 (1982).
78. Cooper, G. M. *Science* **217**, 801-806 (1982).
79. Feinberg, A. P., Vogelstein, B., Droller, M. J., Baylin, S. B. & Nelkin, B. D. *Science* **220**, 1175-1177 (1983).
80. Rubin, H. *Science* **219**, 1170-1171 (1983).
81. Logan, J. & Cairns, J. *Nature* **300**, 103-105 (1982).
82. Rechavi, G., Givol, D. & Canaani, E. *Nature* **300**, 607-611 (1982).
83. Hayward, W. S., Neel, B. G. & Astrin, S. M. *Nature* **290**, 475-480 (1981).
84. Payne, G. S., Bishop, J. M. & Varmus, H. E. *Nature* **295**, 209-214 (1982).
85. Saedler, H., Reif, H. I., Hu, S. & Davidson, N. *Molec. gen. Genet.* **132**, 265-289 (1974).
86. Lautenberger, J. A., Schulz, R. A., Garon, C. F., Tschlis, P. H. & Papas, T. S. *Proc. natn. Acad. Sci. U.S.A.* **78**, 1518-1522 (1981).
87. Quade, K. *Virology* **98**, 461-465 (1979).
88. Cooper, G. A. & Neiman, P. E. *Nature* **292**, 857-858 (1981).
89. Goubin, G., Goldman, D. S., Luce, J., Neiman, P. E. & Cooper, G. M. *Nature* **302**, 114-119 (1983).
90. Tschlis, P. N., Strauss, P. G. & Hu, L. F. *Nature* **302**, 445-449 (1983).
91. Yoshimura, F. K. & Levine, K. L. *J. Virol.* **45**, 576-584 (1983).
92. Kettmann, R. *et al. Proc. natn. Acad. Sci. U.S.A.* **79**, 2465-2469 (1982).
93. Westin, E. H. *et al. Proc. natn. Acad. Sci. U.S.A.* **79**, 2490-2494 (1982).
94. Scolnick, E. M., Weeks, M. O., Shih, T. Y., Ruscetti, S. K. & Dexter, T. M. *Molec. cell. Biol.* **1**, 66-74 (1981).
95. Eva, A. *et al. Nature* **295**, 116-119 (1982).
96. Cairns, J. *Nature* **289**, 353-357 (1981).
97. Klein, G. *Nature* **294**, 313-318 (1981).
98. Marcu, K. B. *et al. Proc. natn. Acad. Sci. U.S.A.* **80**, 519-523 (1983).
99. Klein, G. *Cell* **32**, 311-315 (1983).
100. Adams, J. M. *et al. Proc. natn. Acad. Sci. U.S.A.* **79**, 6966-6970 (1982).
101. Shen-Ong, G. L. C., Keath, E. J., Piccoli, S. P. & Cole, M. D. *Cell* **31**, 443-452 (1982).
102. Taub, R. *et al. Proc. natn. Acad. Sci. U.S.A.* **79**, 7837-7841 (1982).
103. Dalla-Favera, R., Martinotti, S. & Gallo, R. C. *Science* **219**, 963-967 (1983).
104. Erikson, J., ar-Rushidi, A., Drwings, H. L., Nowell, P. C. & Croce, C. M. *Proc. natn. Acad. Sci. U.S.A.* **80**, 820-824 (1983).
105. Crews, S., Barth, R., Hood, L., Prehn, J. & Calame, K. *Science* **218**, 1319-1321 (1982).
106. Rowley, J. D. *Nature* **301**, 290-291 (1983).
107. Yunis, J. J. *et al. New Engl. J. of Med.* **307**, 1231-1236 (1982).
108. deKlein, A. *et al. Nature* **300**, 765-767 (1982).
109. Fialkow, P. J. *et al. Blood* **58**, 158-163 (1981).
110. Dalla-Favera, R. D., Wong-Staal, F. & Gallo, R. C. *Nature* **299**, 61-63 (1982).
111. Collins, S. & Groudine, M. *Nature* **298**, 679-681 (1982).
112. Alitalo, K., Schwab, M., Lin, C. C., Varmus, H. E. & Bishop, J. M. *Proc. natn. Acad. Sci. U.S.A.* **80**, 1707-1711 (1983).
113. Alt, F. W., Kellems, R. E., Bertino, J. R. & Schimke, R. T. *J. biol. Chem.* **253**, 1357-1370 (1978).
114. Murray, M. J. *et al. Cell* **25**, 355-361 (1981).
115. Schwab, M. *et al. Nature* **303**, 497-501 (1983).
116. Chattopadhyay, S. K. *et al. Nature* **296**, 361-363 (1982).
117. Knudson, A. G. *Cancer* Vol. 1 (ed. Becker, F.) 73-88 (Plenum, New York, 1981).
118. Berenblum, I. *Cancer* Vol. 1 (ed. Becker, F.) 451-484 (Plenum, New York, 1981).
119. Bister, K. & Duesberg, P. H. *Proc. natn. Acad. Sci. U.S.A.* **76**, 5023-5027 (1979).
120. Frykberg, L. *et al. Cell* **32**, 227-238 (1983).
121. Balmain, A. & Pragnell, I. B. *Nature* **303**, 72-74 (1983).

ARTICLES

Nd and Sr isotopic study of a mafic layer from Ronda ultramafic complex

A. Zindler*, H. Staudigel*, S. R. Hart†, R. Endres & S. Goldstein*

* Lamont-Doherty Geological Observatory, Department of Geological Sciences, Columbia University, Palisades, New York 10964, USA

† Center for Geoalchemy, Department of Earth and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

Detailed isotopic and trace element analyses of a garnet clinopyroxenite layer from the Ronda ultramafic complex demonstrate the difficulty in obtaining primary mantle information from orogenic lherzolites. However, isotope systematics suggest that mafic layers at Ronda were formed 22 ± 2 Myr ago from a reservoir with initial Nd and Sr isotope ratios of 0.51317 and 0.7025, respectively, and represent the products of interaction between normal mid-oceanic ridge tholeiite and depleted mantle peridotite.

QUESTIONS concerning the nature of the Earth's upper mantle have dominated research efforts in trace element and isotope geochemistry over the past decade. Until recently, however, most geochemical information bearing on the nature of the upper mantle has been obtained indirectly, through studies of basalts and other volcanic rocks erupted at the Earth's surface. While we have learned a great deal from these studies, the

information which derives from them is necessarily model-dependent. Thus, we can only infer mantle characteristics through studies of basalts to the extent that we understand details of magma generation in the mantle.

In recent years, geochemists have turned to ultramafic rocks of presumed mantle origin in their efforts to characterize further the upper mantle. These ultramafic rocks have three principal

modes of occurrence at the Earth's surface: (1) discrete nodules of peridotite and eclogite, ranging from millimetres to metres in dimension, which are brought to the surface by alkalic volcanoes and kimberlites; (2) the lower portions of some ophiolite complexes (such as Bay of Islands, Troodos, and Oman), typically comprising highly-sheared harzburgites; and (3) great masses of peridotite (dunite to aluminous lherzolite), ranging from tens of square metres to hundreds of square kilometres in extent, which have been tectonically emplaced into the crust. These great peridotite massifs, which have been variously called alpine peridotite, orogenic lherzolite, or high-temperature peridotite, are perhaps the most attractive for geochemical study for they do not suffer from the context and scale problems inherent in nodules, nor do they display the nearly total depletion in large-ion lithophile (LIL) elements observed in ophiolite harzburgites. Unfortunately, the secrets held by these rocks have largely eluded geochemists because of a combination of very low trace-element abundances and their typically altered character (the high-temperature, high-pressure, mafic and ultramafic phases are extremely susceptible to alteration by the various processes which act to degrade rocks in the Earth's crust). The success of future geochemical studies of these mantle-derived bodies will depend largely on the geochemists' ability to see through the effects of alteration and decipher the primary mantle chemical code still present in these rocks.

Recent isotopic investigations of orogenic lherzolites¹⁻³ have demonstrated that these bodies are markedly heterogeneous, although there is no systematic study of the scale length of these heterogeneities. Furthermore, the observed isotopic variability has effectively precluded the positive identification of a particular mantle tectonic setting as the source region for these peridotite bodies. Important conclusions based on these recent studies may be summarized as follows: (1) $^{87}\text{Sr}/^{86}\text{Sr}$ ratios for acid-washed clinopyroxene mineral separates tend to lie within the range observed for 'normal' and 'enriched' mid-oceanic ridge basalts (MORBs) (0.7020–0.7035); (2) $^{87}\text{Sr}/^{86}\text{Sr}$ ratios for acid-washed whole-rock lherzolites, harzburgites and clinopyroxenites extend from values typical of MORBs to values in excess of 0.710; (3) intermineral disequilibrium of Sr isotope ratios is typically observed but meaningful isochrons do not emerge; and (4) $^{143}\text{Nd}/^{144}\text{Nd}$ ratios in whole rocks show a large variation (0.5127–0.5136) compared with that of MORBs (0.5130–0.5133). It has been generally stated that contamination by crustal or seawater Sr may in part be responsible for the high $^{87}\text{Sr}/^{86}\text{Sr}$ ratios measured in some whole rocks and minerals.

The present study was undertaken to develop a more quantitative picture of the contamination noted by previous workers, and in the case of $^{143}\text{Nd}/^{144}\text{Nd}$ ratios, to determine the extent to which isotopic variability is due to the *in situ* decay of the radioactive parent nuclide. We considered that this was an essential step to take at a time when geochemists must decide whether further studies of orogenic lherzolites will be fruitful in terms of more detailed characterizations of mantle material and the chemically active processes which operate within the mantle.

For these purposes, we performed a detailed geochemical and isotopic characterization of a garnet clinopyroxene layer which occurs within the spinel lherzolite zone of the Ronda ultramafic complex. A mafic layer, rather than a peridotite, was chosen for the initial phase of this investigation because the mafic layers are generally better preserved and richer in trace elements than the surrounding peridotite.

Geological background

The Ronda ultramafic complex⁴⁻⁹ outcrops over an area of about 300 km² near the south coast of Spain (Fig. 1a). It is an allochthonous slab, ~1.5 km thick, emplaced into the crust at high temperature (~800 °C) ~22 Myr ago (based on Rb–Sr

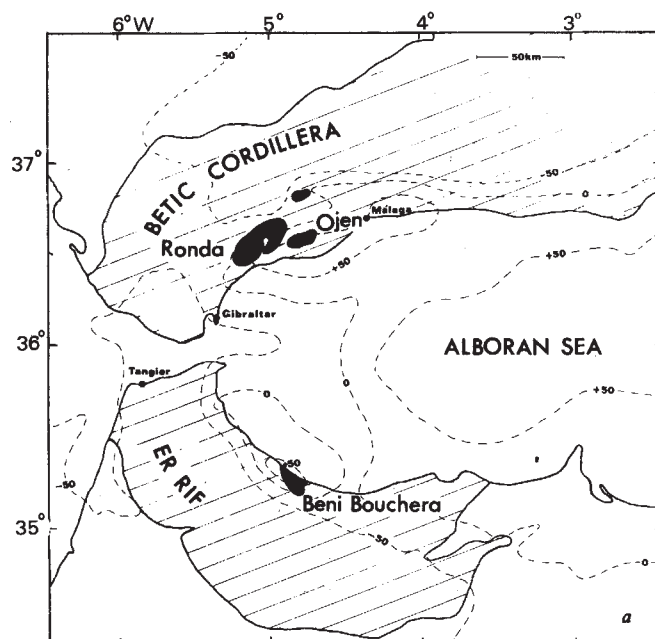
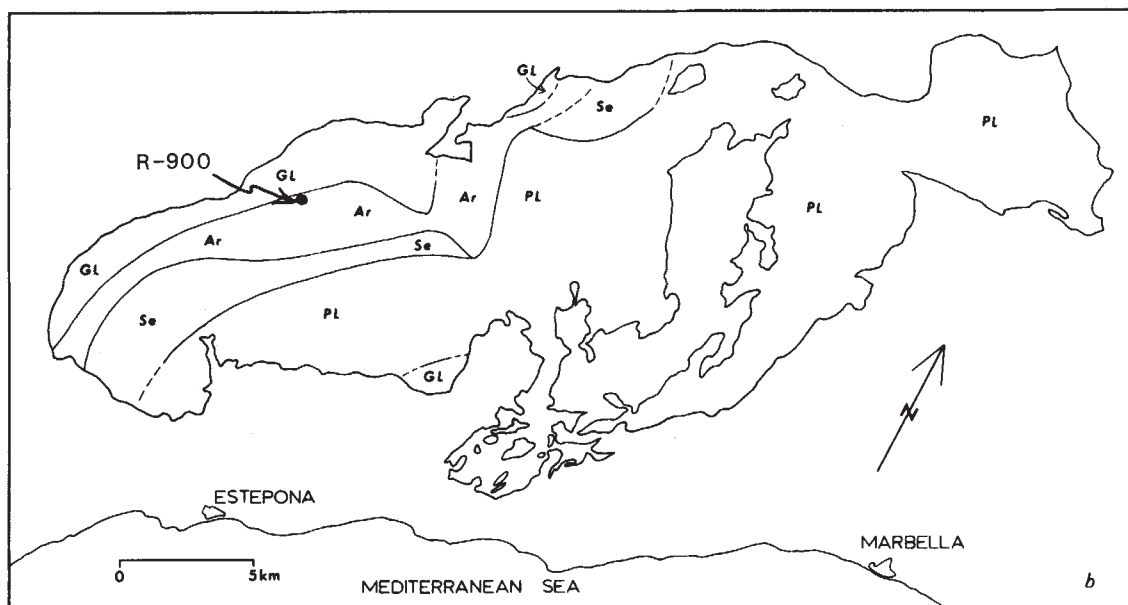


Fig. 1 a, Location map for the Ronda ultramafic complex, taken from ref. 7. Ronda is located at the westernmost end of the Mediterranean Sea in southern Spain. Other peridotite massifs in the Betic-Rif orocline are shown. Bouguer gravity anomalies (in mGal), shown as dashed lines, are from ref. 14. b, Distribution of mineral facies in the Ronda ultramafic complex (taken from ref. 7). GL, garnet lherzolite facies. AR, ariegite subfacies of the spinel lherzolite facies. SE, seiland subfacies of the spinel lherzolite facies. PL, plagioclase lherzolite facies. The location of R-900 is shown.



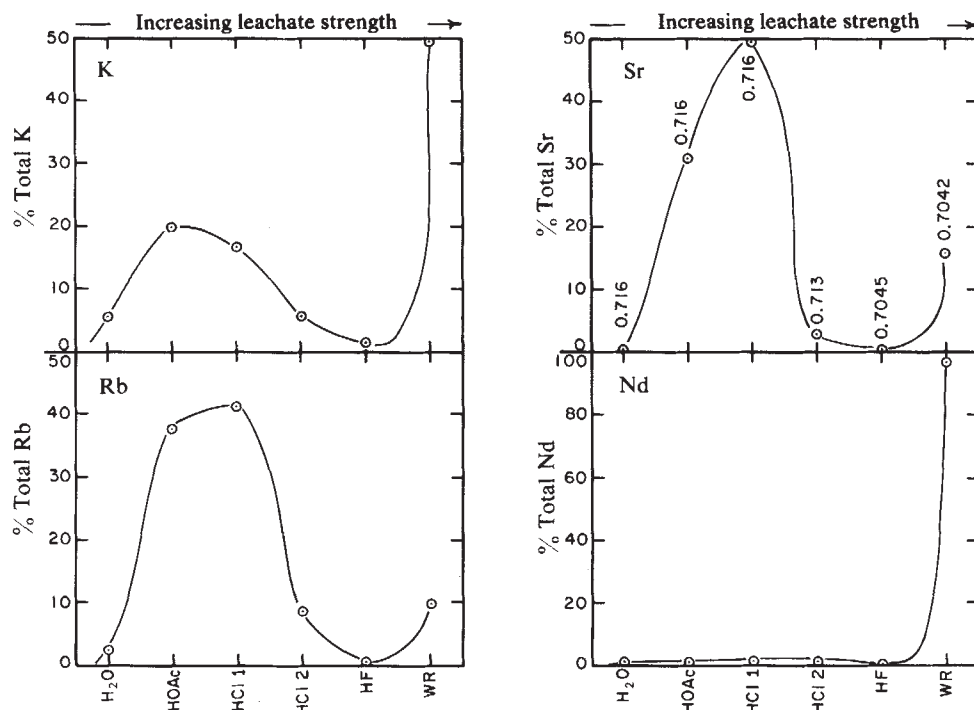


Fig. 2 Mass balance diagram showing the percentage of the total amount of a given element (K, Rb, Sr, Nd) which was found to reside in each of the leachates and in the leached whole rock.

dating of associated anatectic granites¹⁰). The complex contains ~5% mafic 'segregations' ranging in size from several centimetres to several metres and ranging in mineralogy from garnet pyroxenite to olivine gabbro. The peridotite itself shows compositional layering, defined by pyroxene-rich bands up to some tens of centimetres in thickness and hundreds of metres long. In addition, the mineralogy of both the peridotite and mafic layers varies systematically across the massif, and can be used to subdivide the complex into plagioclase, spinel, and garnet lherzolite facies (see Fig. 1b).

The mafic layers, once thought to be more or less pure crystallized liquids derived from partial melting of the peridotites, are now believed to be largely cumulus materials derived from the interaction, at high pressures, of crystallizing partial melts with the surrounding peridotite (depleted magmatic layers of Dickey *et al.*⁷ and Loubet *et al.*¹³). However, this model for the mafic layers, based on rare earth elements (REE) in the peridotites and mafic layers, rests heavily on the assumption that the original peridotite which was partially melted was chemically homogeneous with respect to REE. All of the peridotites and mafic layers so far analysed show light REE depletion (and some depletion in other LIL elements such as K) and appear to be 'residual' in character, the hypothesis being that some fraction of the liquids generated by partial melting has escaped from the massif entirely. Alternatively, it may be argued that the peridotites were residual (depleted) before the melting event which produced the mafic layers. In this case, it appears possible that some of the mafic layers are

close to 'liquid' compositions and indeed may represent 'orthomagmatic layers'⁷.

Analytical methods

The garnet pyroxenite layer chosen for study (R-900) is located in the northwestern section of Ronda within the ariegite sub-facies of the spinel lherzolite facies (see Fig. 1b). R-900 was taken from a 60-cm mafic layer which was investigated by Obata⁵ for melting relations at high pressures.

The experiment was begun with 83 g of sample, free from weathered surfaces and sawmarks and washed ultrasonically in acetone + ethanol and two-bottle distilled water (all reagents used in this experiment were distilled first in a sub-boiling Vycor still, then in a two-bottle Teflon sub-boiling still). The sample was then crushed to the point where all of the material could be passed through a 60-mesh sieve. A series of leaching experiments were then performed: the material was placed in a 1-litre Teflon beaker and about 300 ml of reagent was added; the beaker was placed in an ultrasonic bath under a heat lamp for 20 min; the leachate was then centrifuged until free of particulate matter and decanted; the residue was then further washed with water, centrifuged, and the water added to the leachate; the leachate was evaporated and saved for later analysis and the residue was dried and weighed to determine weight loss during leaching; subsequent leachings were performed in the same manner. After completion of the leaching experiments, a statistically representative aliquot of the solid was separated and used for analysis of a 'leached whole rock'. The very fine

Table 1 Results from leaching experiments on R-900 (concentrations in p.p.m.)

Sample	Sample wt (g)	K	Rb	Cs	Sr	Sm	Nd	K/Rb	Rb/Sr	Sm/Nd	⁸⁷ Sr/ ⁸⁶ Sr
WR (not leached)		60.0	1.66	0.14	34.89	1.810	1.456	36.1	0.0476	1.243	0.70899 ± 4
H ₂ O leach	0.0064	33,913	450	13	1904	1.840	2.437	75.4	0.236	0.7550	0.71564 ± 5
Acetic acid leach	0.89	895	47.8	—	1762	0.716	0.647	18.7	0.0271	1.107	0.71619 ± 3
2.5 M HCl leach (1)	4.13	163	11.1	—	619	—	0.3084	14.7	0.0179	—	0.71587 ± 6
2.5 M HCl leach (2)	2.57	88.7	3.61	0.94	53.3	—	0.3857	24.6	0.0677	—	0.71308 ± 4
5% HF leach	4.11	12.98	0.00814	0.00039	1.657	0.0190	0.02250	1595	0.0491	0.844	0.70449 ± 7
WR (leached)	71.578	29.1	0.152	0.050	11.76	1.441	1.081	191	0.0129	1.333	0.70423 ± 4
Calculated											
unleached WR	83.284	49	1.34	—	61	—	0.96	36.6	0.0220	—	0.71398
Calculated WR											
Measured WR		0.81	0.81	—	1.75	—	0.66	—	—	—	—

⁸⁷Sr/⁸⁶Sr values are relative to 0.70800 for Eimer and Amend SrCO₃. Errors represent 2σ_m.

Table 2 Sm–Nd systematics and modes for R-900

Sample	Mode*	Mode†	Sm (p.p.m.)‡	Nd (p.p.m.)‡	$^{147}\text{Sm}/^{144}\text{Nd}$	$^{143}\text{Nd}/^{144}\text{Nd}$ §
WR (leached)			1.441	1.081	0.8060	0.513276 ± 18
Clinopyroxene	0.379	0.368	2.803	2.620	0.6468	0.513251 ± 11
Garnet	0.383	0.365	1.057	0.1980	3.228	0.513629 ± 27
Plagioclase	0.238	0.267	0.0843	0.1624	0.3138	0.513234 ± 34

* Point-count mode based on 5,000 points. Clinopyroxene garnet and plagioclase are normalized to 1.000; 0.9% oxide has been excluded for purposes of comparison.

† Mass-balance mode based on Sm and Nd concentrations. This mode is consistent with $^{143}\text{Nd}/^{144}\text{Nd}$ data within stated error limits.

‡ Blanks for Sm and Nd are 30–60 pcg and 100–200 pcg, respectively.

§ Errors represent $2\sigma_m$. Measured value of $^{143}\text{Nd}/^{144}\text{Nd}$ in BCR-1 = 0.512645 ± 10 .

|| Mineral separates were washed twice in hot 2.5 M HCl for 20 min and once in cold 5% HF for 10 min. Separates were then finely ground in a boron carbide mortar with cold 2.5 HCl.

material (–200 mesh) was then separated from the remaining solid and the coarse fraction used for mineral separations.

Results and discussion

The results obtained thus far fall naturally into two groups: (1) trace element and Sr isotope compositions of leachable fractions (see Table 1 and Fig. 2); and (2) Sm–Nd systematics of separated major phases (see Table 2 and Fig. 3). We shall discuss the results in this order.

$^{87}\text{Sr}/^{86}\text{Sr}$ ratios for the water, acetic, and HCl leaches (Table 1) range from 0.713 to 0.716 and demonstrate the presence of an easily leachable high $^{87}\text{Sr}/^{86}\text{Sr}$ component(s) in R-900. Further, the low K/Rb ratios and high Sr contents suggest that this component contains some fraction of carbonate material. The high $^{87}\text{Sr}/^{86}\text{Sr}$ ratios, well in excess of seawater values over the past 30 Myr demand the involvement of crustal Sr. These results are in qualitative agreement with those presented by Polve and Allègre³ and indicate that during the emplacement of Ronda into the crust, the presence of an active hydrothermal system facilitated chemical exchange between Ronda and the surrounding crustal materials. A very important aspect of these results is that even after the water, acetic and first HCl leaches, there remained an HCl soluble component with an $^{87}\text{Sr}/^{86}\text{Sr}$ ratio of 0.713.

The 5% HF leach, designed to dissolve the surface layers of major silicate minerals, was expected to yield the same isotopic composition as the leached whole rock. In fact, even the HF leach (0.7045) has dissolved some material higher in $^{87}\text{Sr}/^{86}\text{Sr}$ than the average whole rock (0.7042). The only mineral analysed for Sr isotopes is plagioclase with an $^{87}\text{Sr}/^{86}\text{Sr}$ ratio of 0.70250 ± 4 , and a Sr concentration of 63 p.p.m.. In view of the low $^{87}\text{Sr}/^{86}\text{Sr}$ of the plagioclase, the only major phase which contains a significant amount of Sr, we feel that both the HF solution and the leached whole rock still contain some high $^{87}\text{Sr}/^{86}\text{Sr}$ crustal Sr.

The trace-element concentration data and sample weight data from Table 1 have been used to examine mass balance for the various trace elements in R-900. In Fig. 2, we compare the percentages of the total K, Rb, Sr, and Nd which were found to reside in each of the leachates and the leached whole rock. This diagram demonstrates that: a major portion of the whole-rock budget of Rb, Sr, and K (to a lesser extent) is found to reside in easily leachable phases of presumed secondary origin. The high $^{87}\text{Sr}/^{86}\text{Sr}$ ratios in the leachates demonstrates that they do not significantly attack the primary minerals. The results depicted in Fig. 2 leave little doubt as to the difficulty of obtaining meaningful $^{87}\text{Sr}/^{86}\text{Sr}$ ratios and alkali metal abundances from altered mafic and ultramafic rocks.

Differences between the calculated and measured whole rocks (see Table 1), particularly in Sr content and isotopic composition, reflect the patchy distribution of contaminants, which may be concentrated along minute fractures and at grain boundaries, as shown by Polve and Allègre³. It is unfortunate that the unleached whole rock analysis was performed on a small fragment of R-900, rather than on an aliquot of the sample which was later leached. Some fracture filling was ob-

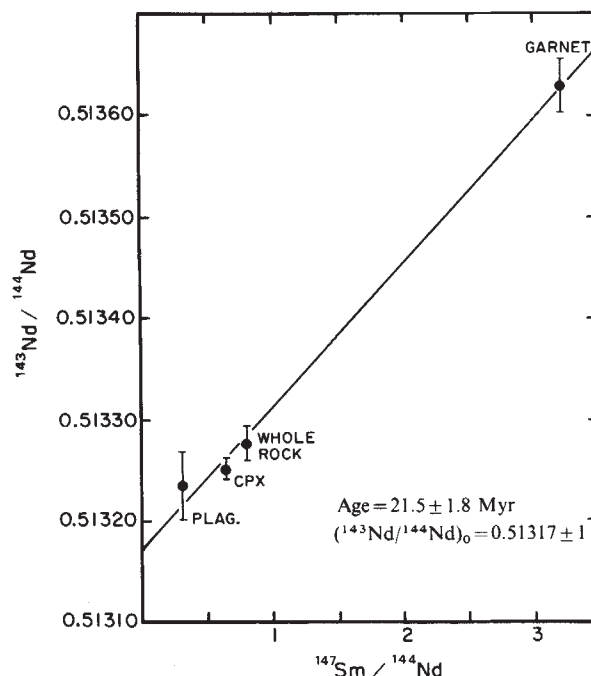


Fig. 3 Sm–Nd isochron diagram for minerals and whole rock from R-900. Regression analysis yields an age of 21.5 ± 1.8 Myr and an initial $^{143}\text{Nd}/^{144}\text{Nd}$ ratio of 0.51317 ± 1 ($\epsilon_{\text{Nd}} = 11$).

served in one thin section, but as yet we have not been able to isolate or determine the composition of this material.

The results from these leaching experiments indicate that Ronda has suffered extensive contamination by crustal materials. Although we have worked on only one sample from Ronda, we suspect that whole-rock Sr isotope work from orogenic lherzolites in general may be perturbed whether or not samples were washed in acid. Further, high $^{87}\text{Sr}/^{86}\text{Sr}$ ratios measured in olivine and orthopyroxene separates by Menzies and Murthy¹ and Polve and Allègre³ must also be regarded with suspicion, as pointed out by these authors. The only clearly meaningful Sr isotope data from orogenic lherzolites would seem to be those analyses of diopsides from Menzies and Murthy¹ and Polve and Allègre³. As discussed earlier, the diopside values span the range observed for 'normal' and 'enriched' modern MORBs (0.7020–0.7035), and indicate that heterogeneities of this order of magnitude exist within orogenic lherzolite bodies.

Although $^{143}\text{Nd}/^{144}\text{Nd}$ ratios were not measured in the leachates, Sm and Nd concentrations in Table 1, and the Nd mass balance in Fig. 2, show that REE are present in the leachable component. Furthermore, the Sm/Nd ratios for the leachates, which are lower than the whole rocks, allow the possibility that the leachable Nd was of crustal origin with low $^{143}\text{Nd}/^{144}\text{Nd}$. Thus, while $^{143}\text{Nd}/^{144}\text{Nd}$ analyses of unleached whole rocks do not seem advisable, the problems presented by

the REE are not nearly as severe as those presented by the alkali metals and alkaline earths. This is further demonstrated by comparison of the calculated and measured modes presented in Table 2.

The second phase of our study of R-900 has been to investigate the nature of Sm–Nd systematics in purified, acid-washed mineral separates and the leached whole rock. The results are listed in Table 2 and are plotted on a Sm–Nd isochron diagram in Fig. 3. Figure 3 shows that minerals are not currently in equilibrium, and that an isochronous relationship does exist, suggesting an age of 21.5 ± 1.8 Myr with an initial $^{143}\text{Nd}/^{144}\text{Nd}$ ratio of 0.51317 ± 1 . Due to great uncertainties in the pressure–temperature–time history of Ronda, we cannot be certain what event this age documents. The 22 ± 4 Myr Rb–Sr age of Priem *et al.*¹⁰ documents the anatexis associated with the emplacement of Ronda into the crust. The combined uncertainties from the Rb–Sr age and the Sm–Nd age, when compared with plate tectonic velocities, are sufficient to permit the possibility that Ronda was deep within the upper mantle when the Sm–Nd clock was activated. Thus, the age could document the formation of the mafic layer (through reaction of a basaltic liquid with the surrounding peridotite), or it could simply document the time when Ronda cooled past the temperature of effective intermineral diffusive equilibration of Nd isotopes. We expect to distinguish between these possibilities by making similar measurements on a garnet peridotite. If peridotites and mafic layers from various parts of the massif yield the same Sm–Nd age, the latter interpretation is supported, although the first cannot be categorically excluded. While the results of Richard and Allègre² serve to characterize orogenic lherzolites as having Sr and Nd isotope ratios typical of the oceanic mantle, the initial $^{143}\text{Nd}/^{144}\text{Nd}$ ratio of 0.51317, together with the $^{87}\text{Sr}/^{86}\text{Sr}$ ratio in the plagioclase (0.70250), show R-900 to have an isotopic character which has been previously observed only at 'normal' mid-ocean ridge segments.

It has been noted (see for example, ref. 1) that the typically low K/Rb ratios in orogenic lherzolites suggest affinities with alkali basalts. Because no alkali basalts have yet been reported with Sr and Nd isotopic compositions similar to R-900, one must consider this unlikely. Although details of the origin of mafic layers in orogenic lherzolites remain elusive, there is some suggestion that they form in conjunction with liquids in the mantle. The isotopic character of R-900 indicates that these liquids were mid-ocean ridge tholeiites, perhaps in the early stages of evolution. In this context, the low K/Rb ratio in the leached whole-rock sample of R-900 (K/Rb = 191) may indicate that, even after leaching, the sample retains some fraction of the material which was partly removed during the leaching experiments. Because of the high alkali metal concentrations and very low K/Rb ratios in the leachable materials (see Table 1), a minute amount of this material would be sufficient drastically to lower the K/Rb ratio in the whole rock.

Perhaps the most important result to come from the Sm–Nd phase of this study is that interpretable isotope systematics are observed. We feel that much information regarding the nature of the upper mantle and upper mantle processes remains to be discovered in orogenic lherzolites.

We thank John Dickey for introducing us to the intrigue that is Serrania de la Ronda, and for a helpful review of an early version of this manuscript, also Claude Allègre and Pierre Richard for making their techniques for chemical separation of Sm and Nd available, Teri Leslie for help with chemical preparation of samples, Patty Catanzaro for the graphics, and Joan Totton for manuscript preparation. Bill White, Dave Walker, and Mireille Polve have provided helpful reviews of this manuscript. This work has been supported in part by a grant from the Harry Oscar Wood Fund of the Carnegie Institution of Washington to S.R.H., and by NSF grants EAR 80-11461 and EAR 82-06477 to A.Z. Lamont-Doherty Geological Observatory contribution no. 3489.

Received 13 October 1982; accepted 20 April 1983.

1. Menzies, M. & Murthy, V. R. *Earth planet. Sci. Lett.* **38**, 346 (1978).
2. Richard, P. Allègre, C. J. *Earth planet. Sci. Lett.* **47**, 65 (1980).
3. Polve, M. & Allègre, C. J. *Earth planet. Sci. Lett.* **51**, 71 (1980).
4. Dickey, J. S. *Miner. Soc. Am. Spec. Publ.* No. 3, 33 (1970).
5. Obata, M. *J. Petrol.* **21**, 933 (1980).
6. Suen, C. J. thesis, Massachusetts Institute of Technology (1978).
7. Dickey, J. S., Obata, M. & Suen, C. J. in *Origin and Distribution of the Elements* (ed. Ahrens, L. H.) 587 (1979).

8. Obata, M. & Lundeen, M. T. *Geol. Soc. Am. Map and Chart Ser.* No. MC-29 (1979).
9. Dickey, J. S., Obata, M. & Lundeen, M. T. *Explanatory text to accompany Geol. Soc. Am. Map and Chart Ser.* No. MC-29 (1979).
10. Priem, H. N. A. *et al.*, *Contr. Miner. Petrol.* **70**, 103 (1979).
11. Hart, S. R. & Brooks, C. *Contr. Miner. Petrol.* **61**, 109 (1977).
12. Zindler, A., Hart, S. R., Frey, F. A. & Jakobsson, S. P. *Earth planet. Sci. Lett.* **49**, 249 (1979).
13. Loubet, M., Polve, M., Richard, P. & Allègre, C. J. in *Colloq. int. Cent. natn. Rech. Scient.* **272**, 269 (1980).
14. Bonini, W. E., Loomis, T. P. & Robertson, J. D. *J. geophys. Res.* **78**, 1372 (1973).

α_1 -Antitrypsin deficiency detection by direct analysis of the mutation in the gene

Vincent J. Kidd, R. Bruce Wallace*, Keiichi Itakura* & Savio L. C. Woo[†]

Howard Hughes Medical Institute, Department of Cell Biology, Baylor College of Medicine, Houston, Texas 77030, USA

* Division of Biology, Molecular Genetics Section, City of Hope Research Institute, Duarte, California 91010, USA

A deficiency in the plasma protease inhibitor α_1 -antitrypsin can cause chronic obstructive emphysema or infantile liver cirrhosis. This deficiency results from a single amino acid substitution created by a G to A transition in the gene for α_1 -antitrypsin. Chemically synthesized specific oligonucleotide probes (19-mer) have been used to develop a sensitive and direct test for the presence or absence of the mutant gene in any individual, which can be used for prenatal diagnosis of the deficiency syndrome.

α_1 -ANTITRYPSIN is an important plasma protease inhibitor capable of inhibiting a wide variety of serine proteases¹. It is synthesized in the liver¹, but a primary site of its physiological action is in the alveolar structure of the lung². α_1 -Antitrypsin acts in the lung by inhibiting polymorphonuclear leukocyte elastase, protecting the elastic fibres of this organ from hydroly-

sis³⁻⁵. When an imbalance between the levels of α_1 -antitrypsin and elastase exists such that the latter is in excess, permanent damage to the alveolar structure of the lung will occur, resulting in eventual development of pulmonary emphysema⁶. Consequently, individuals suffering from a genetic deficiency of α_1 -antitrypsin are predisposed to degenerative lung disease with a risk 20–30 times that of the general population⁷. The severity of this degenerative lung disease in deficient individuals

[†] To whom reprint requests should be addressed.

Table 1 Sequencing analysis of the human α_1 -antitrypsin gene

a

Phenotype	Condition	Genotype
M	Normal	. . .Ala val leu thr ile asp ³⁴² <u>glu</u> lys gly thr glu ala ala.GCT CTG CTG ACC ATC GAC GAG AAA GGG ACT GAA GCT GCT. . . ↓
Z	Deficient	. . .GCT CTG CTG ACC ATC GAC AAG AAA GGG ACT GAA GCT GCT.Ala val leu thr ile asp <u>lys</u> lys gly thr glu ala ala. . .

b

M Coding sequence	5'...ACC ATC GAC GAG AAA GGG A...3'
M-Specific oligonucleotide	3'...TGG TAG CTG CTC TTT CCC T...5'
Z Coding sequence	5'...ACC ATC GAC AAG AAA GGG A...3'
Z Anticoding sequence	3'...TGG TAG CTG TTC TTT CCC T...5'
Z-Specific oligonucleotide	5'...ACC ATC GAC AAG AAA GGG A...3'
M Anticoding sequence	3'...TGG TAG CTG CTG TTT CCC T...5'

a Shows the amino acid sequence and the corresponding nucleotide sequence at the mutation site in the chromosomal α_1 -antitrypsin gene responsible for the deficiency. The G to A transition responsible for the mutation is indicated by the arrow. b Shows the design of the synthetic oligonucleotide probes. The oligonucleotide sequences are shown in bold type, and the M and Z chromosomal DNA sequences are shown above and below this oligonucleotide sequence. The pertinent single base pair mismatches between the oligonucleotide probe and genomic sequences are boxed.

is directly influenced by environmental factors such as industrial pollutants and cigarette smoking^{8,9}, from which deficient individuals are advised to abstain, to reduce the rate of further lung destruction. In addition, during early infancy approximately 14% of the deficient individuals will develop severe liver cirrhosis which is frequently fatal¹⁰. Families who have had previous children with the liver disease often seek prenatal diagnosis of the deficiency syndrome in subsequent pregnancies through fetoscopy.

The synthesis of α_1 -antitrypsin in the liver is controlled by an autosomal and allelic gene system^{11,12}. The α_1 -antitrypsin locus has been assigned the term Pi (P for protease and i for inhibitor) and each phenotype has been assigned a capital letter. Over 30 different phenotypes of the protein have been classified^{13,14}. Most individuals are homozygous for Pi^M and show the normal phenotype. Individuals with the Pi^Z phenotype suffer from a severe deficiency of α_1 -antitrypsin whereas most of the other phenotypes result in relatively normal levels of α_1 -antitrypsin. The plasma level of α_1 -antitrypsin in homozygous individuals showing the Pi^Z phenotype is only 10–15% that of the Pi^M individuals¹⁴. The genetic deficiency involving the Pi^Z phenotype is inherited through an autosomal recessive trait. It has been estimated that approximately 1/2,000 of caucasians are ZZ homozygotes and 5% of the caucasian population are carriers of the Z gene¹⁵. The plasma level of α_1 -antitrypsin in heterozygous individuals with the MZ genotype is about 50–60% of normal¹², suggesting that the defect resides within the α_1 -antitrypsin gene itself and not through a trans-regulatory mechanism.

Although the phenotype of adults can be readily defined by conventional Pi-typing of plasma α_1 -antitrypsin, prenatal diagnosis of the deficiency syndrome is only possible by analysis of fetal blood after fetoscopy, which carries a mortality rate of about 5%. We report here the development of methodologies for prenatal diagnosis of ZZ homozygotes by analysis of the α_1 -antitrypsin gene in amniotic cells which will provide an alternative to the significant risk of fetoscopy.

Mutational sequence

The basis of α_1 -antitrypsin deficiency is the substitution of the glutamic acid residue at position 52 from the carboxyl-terminus

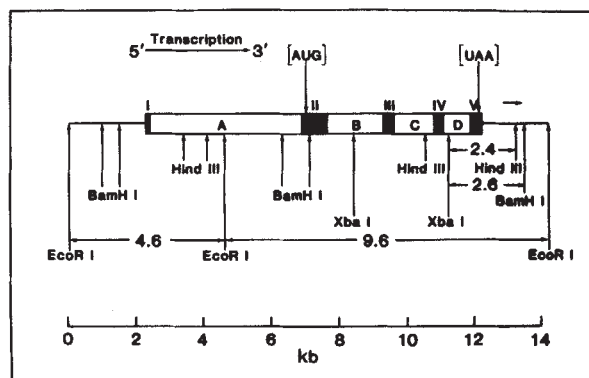


Fig. 1 Structure of the human chromosomal α_1 -antitrypsin gene. The various structural segments (■) are indicated by roman numerals and the intervening sequences (□) by capital letters. The transcriptional orientation of the gene is from left to right, as indicated by the arrow. The mutation site responsible for the deficiency syndrome is found in exon V. Two restriction fragments that contain the mutation site are indicated by the arrows below exon V.

of the normal M-type protein by a lysine residue in the Z-type protein^{16,17}. We have previously reported the construction of α_1 -antitrypsin cDNA clones from baboon and human and demonstrated that mature α_1 -antitrypsin is comprised of 394 amino acids¹⁸. Sequencing analysis of the human complementary DNA (cDNA) clone¹⁸ has demonstrated that the codon for the glutamic acid residue at position 342 of the protein is GAG (Table 1). Since there are only two possible codons for lysine (AAA or AAG), it is very likely that the α_1 -antitrypsin deficiency is caused by a point mutation involving a G to A transition at the corresponding position in the gene (Table 1a).

Studies have shown that synthetic oligonucleotides hybridize specifically to cDNA sequences^{19–21}. Indeed, a single base pair mismatch produces dramatic destabilization in the duplex structure between the synthetic oligonucleotide and the complementary DNA sequence²⁰. Recently, Conner *et al.*²² have shown that the single base pair change in the β -globin gene responsible for the sickle cell disease can be detected in genomic DNA of affected individuals or trait carriers by using synthetic oligonucleotides. Thus, specific oligonucleotides complementary to the M and Z genotypes of the α_1 -antitrypsin gene were synthesized in order to establish a molecular method for the prenatal distinction of these genes.

Synthetic probes

The M- and Z-specific oligonucleotides were designed using the criteria of oligonucleotide length, the position of the mutation site within the probe and the nature of the mismatch formed between the synthetic probe and the genomic DNA (Table 1b). It has previously been shown that synthetic oligonucleotides of 19 residues in length (19-mer) have a high probability of recognizing a unique sequence in the human genome²². By designing the oligonucleotide probes such that there are 9 nucleotides on either side of the mutation site, the thermal instability of mismatched hybrids would be maximized²⁰. For the M-specific probe, a 19-long oligonucleotide complementary to the coding sequence was synthesized. This design produced a C residue at the mutation site which would form a stable C·G base pair between the oligonucleotide probe and the M-coding genomic DNA sequence (Table 1b). This M-specific complementary oligonucleotide sequence forms a C·A mismatch at the mutation site in the Z-coding genomic DNA sequence. The Z-specific oligonucleotide was designed in a similar manner, and in this case the coding sequence of the Z gene was synthesized (Table 1b). At the mutation site the Z-specific probe contains an adenine which would form a stable A·T base pair with the Z-complementary DNA sequence and an A·C mismatch with the M-complementary DNA sequence (Table 1b). The C·A

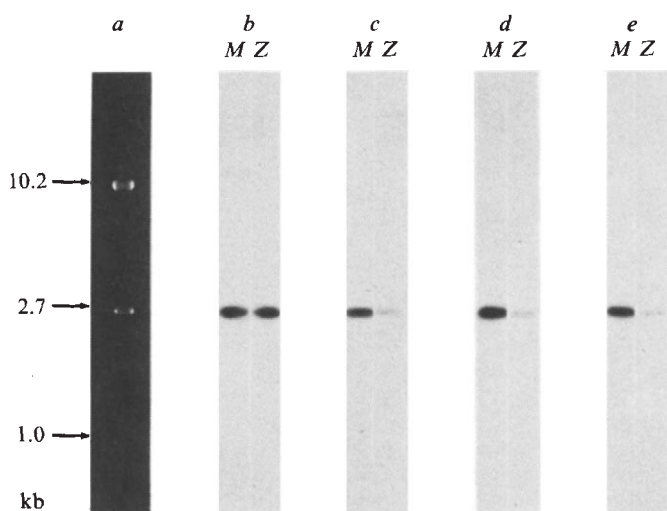


Fig. 2 Selectivity of oligonucleotide hybridization to temperature using gel electrophoresis. *a*, The ethidium bromide stain of the gel; *b*, hybridization with either the *M*- or *Z*-specific oligonucleotide probe at 45 °C and washing at 4 °C; *c*, hybridization with the probes at 45 °C and washing at 4 °C with an additional 1-min wash at 55 °C; *d*, hybridization with the probes at 55 °C and washing at 4 °C; *e*, hybridization with the probes at 55 °C and washing at 4 °C with an additional 1-min wash at 55 °C. Nonadecanucleotides corresponding to the *M* and *Z* genomic sequences (Table 1) were synthesized on a solid support by the modified triester approach, as described by Itakura *et al.*^{26,27}, and labelled with [γ -³²P]ATP (ICN, crude preparation, >7,000 Ci mmol⁻¹) by a kinase reaction¹⁹ to a specific activity of at least 3×10^8 c.p.m. μ g. The labelled oligonucleotide was separated from unlabelled nonadecanucleotide and reaction products by electrophoresis on a 20% polyacrylamide-7 M urea denaturing gel and the labelled oligonucleotide was eluted by diffusion. The pAT9.6 plasmid²³ contains a 9.6-kb *Eco*RI fragment of the normal human α_1 -antitrypsin gene. This plasmid DNA was digested according to manufacturer's specifications with *Hind*III endonuclease (BRL). The agarose was filtered through a 0.8- μ m Millipore nitrocellulose filter after boiling. The DNA (0.2 μ g) was loaded in duplicate lanes of a horizontal (2 mm thick) 1.0% agarose gel (Seakem) and electrophoresed at 1 V cm⁻¹ at 50 V for 750 V-h in a Tris/borate electrophoresis buffer²². The gel was trimmed and then stained with ethidium bromide (Calbiochem) and photographed over UV light with a Polaroid land camera. The DNA was denatured *in situ* by treating with 0.5 M NaOH, 0.15 M NaCl at 4 °C for 30 min while gently shaking and then neutralized with 0.5 M Tris-HCl, pH 8.0, 0.15 M NaCl at 4 °C for 30 min while shaking gently. The gel was placed on two sheets of Whatman 3MM paper, covered with Saran Wrap and dried under vacuum at 60 °C with a Hoefer gel dryer. The dried gel was rinsed in H₂O to remove the backing paper and then hybridized with either the *M*- or *Z*-specific oligonucleotide (3.0×10^8 c.p.m. per μ g) at 10 ng ml⁻¹ probe in 6 $\times$ NET (NET = 150 mM NaCl, 1 mM EDTA and 15 mM Tris-HCl, pH 7.5) and 0.1% SDS at either 45 or 55 °C for 2 h in Seal-a-meal bags. The gel strips were then washed in four changes of 6 $\times$ SSC, 0.5% SDS for 15 min each at 4 °C. Additionally, some of the gel strips were washed at 55 °C in 6 $\times$ SSC, 0.5% SDS for 1 min. The strips were then wrapped in Saran Wrap and autoradiographed with Kodak XAR-5 X-ray film between two Quanta III intensifier screens (Dupont) at -80 °C for 4 h.

mismatch formed by both the *M* and *Z* synthetic oligonucleotide probes with genomic DNA sequences is preferable to other mismatches (unpublished observations).

Restriction fragments

Restriction fragments 1–3 kilobases (kb) in size have been shown to be optimal for the detection of single point mutations in genomic DNA using synthetic oligonucleotides²². Several restriction fragments of these sizes can be generated from the human chromosomal α_1 -antitrypsin gene so as to contain exon V (Fig. 1)²³ which, by sequence analysis, contains the mutation

responsible for the generation of the *Z* gene (Long *et al.*, unpublished results). These fragments include a 2.7-kb *Hind*III fragment, a 2.6-kb *Bam*HI-*Xba*I fragment and a 2.4-kb *Hind*III-*Xba*I fragment (Fig. 1).

Specificity

The specificity in hybridization of synthetic oligonucleotides to cDNA is controlled by the stringency of the conditions used for hybridization and subsequent washing²¹. These conditions were established by test hybridization of the terminally labelled *M* and *Z* synthetic oligonucleotides to the cloned human α_1 -antitrypsin gene (Fig. 2). In this experiment, pAT9.6 DNA²³, which is a human genomic α_1 -antitrypsin DNA clone containing a GAG at the position corresponding to amino acid residue 342 of the protein, was cleaved with *Hind*III to generate a 2.7-kb fragment containing the mutation site (Fig. 2*a*). Agarose gel strips containing identical *Hind*III-digested pAT9.6 DNA fragments were hybridized with the synthetic *M* and *Z* probes and washed in four different sets of conditions. When hybridization and wash conditions were at low stringency, both the *M* and *Z* probes yielded strong signals with the cloned *M*-type sequence in the gel (Fig. 2*b*). Elevating the temperature for hybridization or washing caused significant reduction of the signal strength generated by the *Z* probe but not the *M* probe, resulting in the generation of differential signals (Fig. 2*c, d*). The maximal differential signal was obtained by elevating the temperature for both hybridization and washing (Fig. 2*e*). Scanning of the X-ray film has indicated that the signal strengths generated by the two probes differed by a minimum of 10-fold (data not shown).

Genotype identification

Using hybridization conditions that yielded maximal differential signals, the synthetic oligonucleotide probes were used to analyse genomic DNAs isolated from fibroblasts of the *MM*, *MZ* and *ZZ* genotypes (from the Human Genetics Mutant Cell Repository). In these experiments, the cloned human α_1 -antitrypsin gene was used as a standard. Multiple DNA fragments were obtained from pAT9.6 DNA when digested with *Hind*III-*Xba*I (Fig. 3*a*, left-hand lane). After digestion, the cloned DNA was serially diluted to contain 10, 5 and 1 copy-equivalents per haploid genome and applied to lanes 1, 2 and 3 of Fig. 3*a*. After hybridization with the *M* probe, signals of corresponding strength were generated by the 2.4-kb DNA fragment which contains exon V of the α_1 -antitrypsin gene, suggesting that the analytical procedure is not only qualitative, but also quantitative. Genomic DNA isolated from fibroblasts of the *MM*, *MZ* and *ZZ* genotypes obtained from the Human Genetic Mutant Cell Depository were then digested with the same enzymes and applied to lanes 5, 6 and 7, respectively, of the same agarose gel (Fig. 3*a*). After hybridization with the synthetic *M* probe, intense signals were generated by multiple DNA bands of higher molecular weight, most probably the result of hybridization of the probe with repetitive DNA sequences in the human genome that are homologous with the probe to various extents. This level of nonspecific hybridization does not affect the analysis because only hybridization signals with the 2.4-kb band are relevant. It is apparent that only DNA of the *MM* genotype in lane 5 produced a strong signal at 2.4 kb, and the intensity of this signal was comparable with that generated from the single copy-equivalent per haploid genome of the cloned *M*-type α_1 -antitrypsin gene sequence (Fig. 3*a*, lane 3). The *MZ* genotypic DNA produced a signal of intermediate strength (Fig. 3*a*, lane 6) and DNA of the *ZZ* genotype showed no hybridization signal at 2.4 kb at all (Fig. 3*a*, lane 7).

The experiment described above was repeated using a *Bam*HI-*Xba*I double digestion of the cloned and genomic DNA samples. The 2.6-kb *Bam*HI-*Xba*I restriction DNA fragment containing exon V of the cloned α_1 -antitrypsin gene also produced signals of corresponding strength when hybridized with the synthetic *M* probe (Fig. 3*b*, lanes 1–3). Once again, *MM*

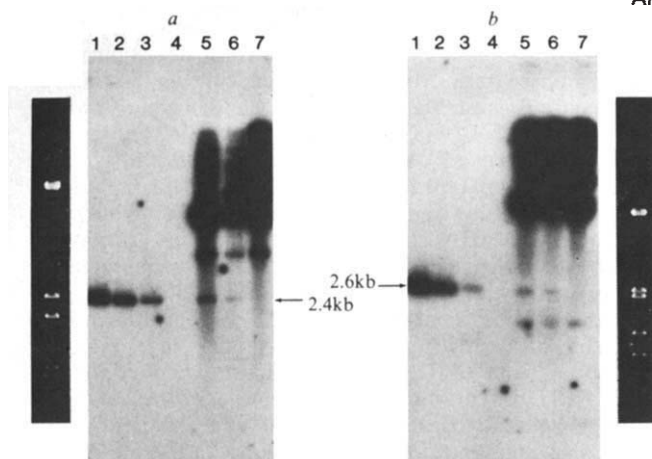


Fig. 3 Hybridization of ^{32}P -labelled *M*-specific oligonucleotide to human genomic DNA digested with different restriction endonucleases. In *a*, an ethidium bromide-stained gel showing the cloned α_1 -antitrypsin gene digested with *HindIII*-*XbaI* is shown on the left. Lanes 1, 2 and 3 are serially diluted pAT9.6 DNA digested with *HindIII*-*XbaI* containing 10, 5 and 1 genome equivalents of the cloned α_1 -antitrypsin gene corresponding to the *MM* genotype; lane 4 contains pBR322 DNA; lane 5, 10 μg of *MM* DNA digested with *HindIII*-*XbaI*; lane 6, 10 μg of *MZ* DNA digested with *HindIII*-*XbaI*; and lane 7, 10 μg of *ZZ* DNA digested with *HindIII*-*XbaI*. In *b*, the same DNAs were digested with *BamHI*-*XbaI* and loaded on the gel in the same order. An ethidium bromide-stained gel showing the cloned α_1 -antitrypsin gene digested with *BamHI*-*XbaI* is shown on the right. **Methods:** Cell lines corresponding to apparently normal fibroblasts (GM 0970), *Pi* type *MZ* lymphoblasts (GM 3579) and *Pi* type *ZZ* lymphoblasts (GM 3578) from the Human Genetic Mutant Cell Repository were isolated as described by Kunkel *et al.*²⁸. Human DNA (10 μg) was digested at 5 enzyme units per μg of DNA at 37°C for 10–12 h. The restriction endonucleases used were *BamHI*-*XbaI*, or *HindIII*-*XbaI* (BRL) and the digests were carried out according to manufacturers' specifications. The digested DNA was run on a 1% agarose gel and the gel processed as described in Fig. 2 legend. The dried gel was hybridized with ^{32}P -labelled *M*-specific oligonucleotide (3.0×10^8 c.p.m. per μg) at 10 ng ml⁻¹ probe for 2 h at 55°C. It was then washed four times in 6×SSC, 0.5% SDS at 4°C for 15 min each, then given a 1-min wash in 6×SSC, 0.5% SDS at 55°C followed by a 1-h wash in 6×SSC, 0.5% SDS at 22°C. The gel was then autoradiographed between two Quanta III intensifier screens (DuPont) for 4 days.

DNA produced a strong signal (Fig. 3*b*, lane 5) corresponding to that generated by the single copy-equivalent per haploid human genome as shown in lane 3 of the same gel. In addition, the *MZ* DNA produced a signal of intermediate strength (Fig. 3*b*, lane 6) and the *ZZ* individual's DNA produced no detectable signal at 2.6 kb (lane 7). Thus, in appropriate conditions the synthetic *M* probe is capable of distinguishing cells of the *MM* and *MZ* genotypes from those of the *ZZ* genotype.

These experiments were repeated using the synthetic *Z* probe (Fig. 4). This time, the lanes containing the serially-diluted cloned *M*-type α_1 -antitrypsin gene sequence showed no hybridization signal (Fig. 4*a*, *b*, lanes 1, 2), again indicating the specificity of the analysis. Only the *ZZ* DNA yielded a strong hybridization signal at 2.4 kb after *HindIII*-*XbaI* digestion (Fig. 4*a*, lane 5), and at 2.6 kb after *BamHI*-*XbaI* digestion (Fig. 4*b*, lane 5). Whereas the *MZ* individuals DNA produced signals of intermediate intensity at these positions (Fig. 4*a*, *b*, lanes 4), the *MM* DNA produced no signal at these positions at all (Fig. 4*a*, *b*, lanes 3). Thus, the synthetic *Z* oligonucleotide probe is also capable of distinguishing *MM* homozygotes and *MZ* heterozygotes from the *ZZ* homozygotes. Use of both probes allows the unambiguous determination of all three genotypes.

Detection of *ZZ* individuals

Although the probes worked well on cells from a repository, it was also of importance to demonstrate that the methodology

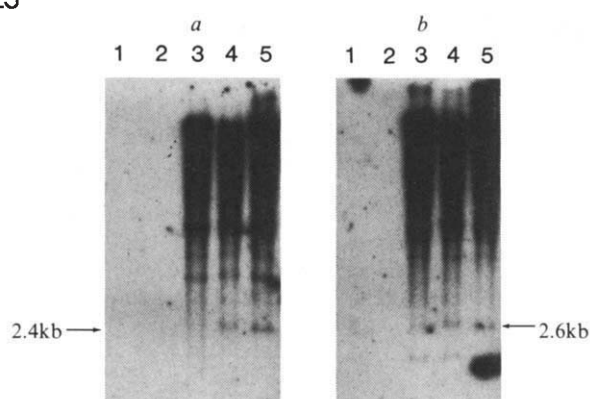


Fig. 4 Hybridization of ^{32}P -*Z*-specific oligonucleotide probe to human genomic DNAs digested with various restriction endonucleases. Chromosomal DNAs from cultured cells of *MM*, *MZ* and *ZZ* genotypes digested with either *HindIII*-*XbaI* (*A*) or *BamHI*-*XbaI* (*B*) were treated as described previously. Lanes 1 and 2 are serially diluted pAT9.6 DNA containing 10 and 5 genome equivalents of the cloned α_1 -antitrypsin gene corresponding to the *MM* genotype; lane 3, 10 μg of *MM* DNA; lane 4, 10 μg of *MZ* DNA; lane 5, 10 μg of *ZZ* DNA.

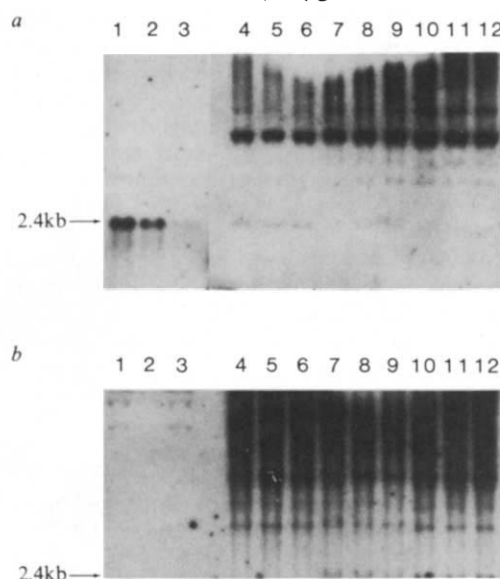


Fig. 5 Detection of *ZZ* individuals. DNA from nine different individuals of various *Pi* types (kindly provided by Drs Jack Pierce and Ronald Crystal) were prepared²⁸ and digested with endonucleases *HindIII*-*XbaI*, run in duplicate on two different 1% agarose gels and processed as described previously. Lanes 1, 2 and 3 contain serially diluted pAT9.6 DNA, digested with *HindIII*-*XbaI*, containing 10, 5 and 1 genome equivalents of the cloned α_1 -antitrypsin gene. Lanes 4–6 contain 10 μg of DNA from three different *MM* individuals. Lanes 7–9 contain 10 μg of DNA from three different *MZ* individuals. Lanes 10–12 contain 10 μg of DNA from three different *ZZ* individuals. In *a*, the dried gel was hybridized with ^{32}P -*M*-specific oligonucleotide probe (3.0×10^8 c.p.m. per μg) at 10 ng ml⁻¹ probe, then washed and autoradiographed as previously described. In *b*, the duplicate dried gel was hybridized with ^{32}P -*Z*-specific oligonucleotide probe (3.0×10^8 c.p.m. per μg) at 10 ng ml⁻¹ probe, then washed and autoradiographed as previously described.

could be satisfactorily applied to cells of different individuals before its application as a diagnostic tool for α_1 -antitrypsin deficiency. Thus, genomic DNA samples were prepared from peripheral lymphocytes of three individuals each of the *MM*, *MZ* and *ZZ* genotypes. The DNA preparations were digested with *HindIII*-*XbaI* and electrophoresed on agarose gels. The gel was hybridized with the synthetic probes and the presence of the 2.4-kb gene-specific bands were revealed by autoradiography. The synthetic *M* probe produced hybridization signals at 2.4 kb in the *MM* homozygotes (Fig. 5*a*, lanes

4-6) and the *MZ* heterozygotes (Fig. 5a, lanes 7-9). More importantly, the *M* probe failed to generate a hybridization band in DNA of all three *ZZ* homozygotes (Fig. 5a, lanes 10-12). Conversely, the *Z*-specific probe produced bands at 2.4 kb in DNA of the three *MZ* heterozygotes (Fig. 5b, lanes 7-9) and the three *ZZ* homozygotes (Fig. 5b, lanes 10-12). Again, the *Z* probe did not hybridize with a 2.4-kb band in DNA samples of the three *MM* homozygotes (Fig. 5b, lanes 4-6). Serially diluted DNA corresponding to 10, 5 and 1 copy-equivalents per haploid genome hybridized with the synthetic *M* probe (Fig. 5a, lanes 1-3) but did not hybridize with the *Z*-specific probe, as expected (Fig. 5b, lanes 1-3). Thus, a combination of *HindIII-XbaI* and *BamHI-XbaI* digestion of genomic DNA samples with the use of the *M*- and *Z*-specific oligonucleotide probes will permit the identification of the genotypes of any given individuals with a rather high level of confidence, and the methodology can be used for prenatal diagnosis of α_1 -antitrypsin deficiency.

Discussion

The probability that a given point mutation will result in a restriction endonuclease recognition site alteration is not great²⁴. However, α_1 -antitrypsin deficiency is the result of such a point mutation. Furthermore, the deficiency cannot be readily diagnosed by restriction fragment length polymorphisms associated with the gene (as has been used in various forms of α - and β -thalassaemia²⁵), due to the lack of enzymes that yield polymorphic patterns at sufficiently high frequencies (unpublished results). In the absence of these established molecular methods for prenatal diagnosis of genetic disorders, we have adapted the recently developed procedure that uses synthetic oligonucleotides to identify β -globin gene of the sickle-cell type, to analyse α -antitrypsin deficiency. The synthetic *M*- and *Z*-specific oligonucleotide probes have allowed us to detect

directly the underlying mutation within the α_1 -antitrypsin gene.

Since only *ZZ* homozygotes are at risk of developing pulmonary emphysema and/or infantile liver cirrhosis, any analytical procedure for prenatal diagnosis of the deficiency must be able to distinguish DNAs of the *ZZ* genotypes from the *MZ* or *MM* genotypes. In controlled hybridization and washing conditions, it is apparent that the probe specific for the normal gene sequence yielded significant hybridization signals only with the chromosomal DNAs of the *MM* and *MZ* genotypes, whereas the probe specific for the deficient gene sequence hybridized only with the chromosomal DNAs of the *MZ* and *ZZ* genotypes. It is remarkable that the oligonucleotide probes are capable of identifying a single nucleotide change among 3×10^9 base pairs of DNA present in the haploid human genome. The use of both oligonucleotide probes and two different enzymatic digests should result in the determination of particular genotypes in any individual with a high level of confidence. Thus, using a combination of this methodology and genetic counselling plus amniocentesis, it will be possible to perform prenatal diagnosis for the deficiency syndrome for families at risk.

Specific oligonucleotide probe analysis of genomic DNA will be applicable to other hereditary disorders caused by point mutations, provided that the nucleotide sequences surrounding the mutation site can be established. This technique has the advantage over analysis by restriction length polymorphism in that it is a direct analysis which does not require the existence of probands.

We thank Dr Jack Pierce and Dr Ronald Crystal for providing us with the lymphocyte samples from individuals of various *Pi* types. This work was supported in part by NIH grant HI 27509 to S.L.C.W. and grant HI 29516 to R.B.W. R.B.W. and K.I. are members of the Cancer Research Center at the City of Hope Research Institute, and S.L.C.W. is an Investigator of the Howard Hughes Medical Institute.

Received 17 March; accepted 24 May 1983.

1. Heimburger, N., Haupt, H. & Schwick, H. G. in *Proc. Int. Res. Conf. Protease Inhibitors* (eds Fritz, H. & Tschesche, H.) 1-22 (de Gruyter, Berlin, 1970).
2. Laurell, C. B. & Erikson, S. *Scand. J. Clin. Lab. Invest.* **15**, 132-140 (1963).
3. Briscoe, W. A., Kueppers, F., Davis, A. L. & Bearn, A. G. *Am. Rev. resp. Dis.* **94**, 529-539 (1966).
4. Sharp, H. L., Bridges, R. A., Krivit, W. & Freier, E. F. *J. Lab. Clin. Med.* **75**, 934-939 (1969).
5. Olsen, G. N., Harris, J. O., Castle, J. R., Waldeman, R. H. & Karmgard, H. J. *J. Clin. Invest.* **55**, 427-430 (1975).
6. Gadek, J. E. *Bull. Eur. Physiopath. Resp.* **16**, 27-40 (1980).
7. Kueppers, F. in *Lung Biology in Health and Disease* Vol. 11 (ed. Litwin, S.) 23 (Dekker, New York, 1978).
8. Gadek, J. E. *Bull. Eur. Physiopath. Resp.* **16**, Suppl., 27-40 (1980).
9. Janus, E. D. & Carrell, R. W. *N. Z. med. J.* **81**, 461-467 (1975).
10. Larsson, C. *Acta med. scand.* **204**, 345-351 (1978).
11. Kueppers, F. *Am. J. hum. Genet.* **25**, 677-686 (1973).
12. Fagerhol, M. K. & Laurell, C. B. *Prog. med. Genet.* **7**, 96-111 (1970).

13. Fagerhol, M. K. *Ser. Haemat.* **1**, 153-161 (1968).
14. Allen, R. C., Harley, R. A. & Talamo, R. C. *Am. J. Clin. Path.* **62**, 732-739 (1974).
15. Laurell, C. B. & Sueger, T. *Am. J. hum. Genet.* **27**, 213-217 (1975).
16. Jeppsson, J. O. *FEBS Lett.* **65**, 195-197 (1976).
17. Yoshida, A., Lieberman, J., Gaidulis, L. & Ewing, C. *Proc. natn. Acad. Sci. U.S.A.* **73**, 1324-1328 (1976).
18. Kurachi, K. *et al. Proc. natn. Acad. Sci. U.S.A.* **78**, 6826-6830 (1981).
19. Wallace, R. B. *et al. Nucleic Acids Res.* **6**, 3543-3557 (1979).
20. Wallace, R. B. *et al. Nucleic Acids Res.* **9**, 879-894 (1981).
21. Wallace, R. B., Schödl, M., Johnson, M. J., Dembek, P. & Itakura, K. *Nucleic Acids Res.* **9**, 3647-3656 (1981).
22. Conner, B. J. *et al. Proc. natn. Acad. Sci. U.S.A.* **80**, 278-282 (1983).
23. Leicht, M. *et al. Nature* **297**, 655-659 (1982).
24. Nei, M. & Li, N. H. *Proc. natn. Acad. Sci. U.S.A.* **76**, 5269-5273 (1979).
25. Orkin, S. H. *et al. Nature* **296**, 627-631 (1982).
26. Miyoshi, K., Huang, T. & Itakura, K. *Nucleic Acids Res.* **8**, 5491-5505 (1979).
27. Dembek, P., Miyoshi, K. & Itakura, K. *J. Am. chem. Soc.* **103**, 706-708 (1981).
28. Kunkel, L. M. *et al. Proc. natn. Acad. Sci. U.S.A.* **74**, 1245-1249 (1978).

Drosophila melanogaster mitochondrial DNA, a novel organization and genetic code

Maarten H. L. de Bruijn*

MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK

The sequence of a 4,869 base-pair fragment of Drosophila melanogaster mitochondrial DNA is presented. It contains genes for cytochrome oxidase subunits I, II and III, ATPase subunit 6 and six tRNAs together with two unassigned reading frames. The gene organization differs from that of mammalian mitochondrial DNAs. Evidence is provided for a genetic code in which AGA codes for serine and the quadruplet ATAA is used in initiation of translation.

MITOCHONDRIAL DNA (mtDNA) of *Drosophila melanogaster* is a covalently closed duplex circle of about 19,500 base pairs (bp)¹. *Drosophila* mtDNA is unusual in having an A + T-rich region which varies in size and sequence between species,

and even between strains, and accounts almost completely for size differences between various *Drosophila* mtDNAs¹⁻⁴. This precludes a serious coding function for the region but it contains an origin from which replication proceeds in one direction⁵.

Two rRNA genes have been located relative to the A + T-rich region (Fig. 2a)⁶. Their tandem arrangement and the direction of transcription from small to large rRNA are as found in

* Present address: Department of Molecular Immunology, Research Institute of Scripps Clinic, La Jolla, California 92037, USA.

Table 1 Characteristics of *D. melanogaster* mitochondrial genes and comparison with the homologous bovine and yeast genes

Gene	Protein length	Molecular weight	% Amino acid conservation		% Nucleotide conservation		% Unchanged codons	
			D/B	D/Y	D/B	D/Y	D/B	D/Y
URF 2*	—	—	35.2	—	49.0	—	18.4	—
CO I	512	56,400	75.8	58.8	70.3	64.2	34.9	33.0
CO II	228	26,200	58.6	46.7	63.0	60.0	26.4	34.3
URF A6L	53	6,300	25.9	—	37.9	—	13.8	—
ATPase 6	224	25,200	39.0	26.9	51.4	43.8	17.1	17.7
CO III*	—	—	59.8	40.1	60.9	52.9	26.3	23.5

The nucleotide and deduced amino acid sequences of the *D. melanogaster* mtDNA fragment (D) were aligned with the homologous bovine (B) and yeast (Y) sequences^{8,23-26}. The main criterion for alignment was amino acid identity, the second amino acid similarity⁶⁰, the third nucleotide identity and the fourth nucleotide similarity (purine or pyrimidine). Where necessary the nucleotide sequences were padded without changing the reading frame (the one exception to this is shown in Fig. 3). Molecular weights and percentages were calculated using the computer programs of Staden (ref. 61 and unpublished results). The percentage of conservative codon changes can be calculated by subtracting the percentage of unchanged codons from the percentage of conserved amino acids. All percentages have been calculated over the length of the shortest of the two sequences compared.

*The gene is only partially encoded in the mtDNA fragment.

mammalian and amphibian mitochondria⁷⁻¹⁰. However, *Drosophila* mtDNAs are unique in that both replication and rRNA transcription proceed in the same direction whereas in vertebrates they proceed in opposite directions⁵.

A number of polyadenylated transcripts, including the large rRNA, have been identified and mapped on *D. melanogaster* mtDNA^{11,12}. They appear to saturate the entire genome outside the A+T-rich region. Recently a number of short nucleotide sequences from *Drosophila yakuba* mtDNA have been reported¹³ which indicate a different gene organization from that of mammalian mtDNAs. Here, the complete sequence of a 4,869-bp fragment of *D. melanogaster* mtDNA is presented. It confirms and extends the differences found in *D. yakuba* and provides evidence for a unique genetic code.

Organization

Figure 1 shows the sequence and organization of the *D. melanogaster* mtDNA fragment. By comparison with bovine mtDNA⁸, genes have been identified for cytochrome oxidase subunits I, II and III (COI-III), ATPase subunit 6 (ATPase 6) and two unassigned reading frames (URFs), URF 2 and URF A6L. The 5' region of URF 2 and the 3' region of CO III are outside this fragment. No significant insertions or deletions were found, nor additional reading frames not represented in mammalian mitochondrial genomes. The results of a detailed alignment of the genes with their bovine and yeast counterparts are shown in Table 1. The sequences differ from yeast genes by insertions and deletions. Six potential tRNA genes were identified using the computer program TRNA¹⁴. According to the mammalian mitochondrial genetic code^{8,15}, the encoded tRNAs specify tryptophan, cysteine, tyrosine, leucine, lysine and aspartic acid.

The genes fit into a pattern of organization similar to that found in mammalian mitochondrial genomes⁷⁻⁹. They are tightly packed, with tRNA genes separating the protein-coding regions. There are few intergenic regions, the largest being a 19-nucleotide sequence between the tRNA^{Cys} and tRNA^{Tyr} genes. All other genes are either buttjointed or separated by at most five nucleotides. The tRNA^{Trp} and tRNA^{Cys} genes overlap by eight nucleotides but are on different strands. As in mammalian mtDNAs the ATPase 6 gene is preceded by a short out-of-phase but overlapping reading frame (URF A6L) on the same strand.

Difference between *Drosophila* and mammals

The sequenced *Hind*III fragment has been positioned and orientated relative to the A+T-rich region on the basis of its size and an asymmetrically positioned *Hae*III recognition site (Fig. 2a). The sense sequence (Fig. 1) and therefore transcription are in an anticlockwise direction. Klukas and Dawid⁶ have shown that transcription of the rRNA genes is from small to large in a clockwise direction. Therefore, rRNA genes and

protein-coding genes are on complementary strands. An identical gene arrangement has been found in *D. yakuba* mtDNA, in which URF 1 was also shown not to precede URF 2 as in mammals but to be on the complementary strand distal to the large rRNA gene¹³.

In Fig. 2b the gene map of the *Drosophila* mtDNA fragment is aligned with the corresponding region in bovine mtDNA⁸. Whereas the protein-coding genes are in equivalent positions the tRNA genes are not. Compared with the bovine genome, the genes for tRNA^{Ala} and tRNA^{Asp} have disappeared together with the adjoining origin of replication. The 19-nucleotide region between the tRNA^{Cys} and tRNA^{Tyr} genes is only about half the size of the mammalian L-strand origin of replication and cannot form an equivalent secondary structure^{8,16}. The tRNA^{Ser}_{UCN} and tRNA^{Asp} genes have been replaced by the tRNA^{Leu}_{UUR} gene, which in mammals is located between the 16S rRNA gene and URF 1. In *Drosophila* the tRNA^{Asp} gene is located 800 nucleotides downstream, between the tRNA^{Lys} gene and URF A6L.

Genetic code

The codon UGA, a termination codon in the 'universal' genetic code, specifies tryptophan in all mammalian and fungal mitochondrial genetic systems studied^{9,15,17-20}. The protein-coding genes in the *D. melanogaster* fragment contain 45 internal TGA codons (Fig. 1), 35 of which align with tryptophan codons in the corresponding bovine sequence. In addition, a potential tRNA^{Trp} with anticodon TCA has been identified (Figs 1, 4b). There is little doubt that in *Drosophila* mitochondria UGA also codes for tryptophan, although alignment with independently determined amino acid sequences will be required to prove this formally.

A second deviation from the 'universal' code concerns the codon AGA which, in that code, is specific for arginine. Both AGA and AGG have been shown to be termination codons in mammalian mitochondrial systems^{7,8} but arginine codons in fungal mitochondria²⁰⁻²². Clary *et al.*¹³ assume that in *D. yakuba* mtDNA the 'universal' code is used. The *D. melanogaster* mtDNA sequence presented here strongly suggests that AGA codes for serine. It occurs 23 times as an internal codon and in all six identified reading frames (Fig. 1), and therefore is not a termination codon. When the *D. melanogaster* genes are compared with their bovine and yeast counterparts^{8,23-26}, none of the AGA codons align with arginine. As shown in Table 2, assignment of AGA to serine appears most likely.

Without independently determined amino acid sequences, this assignment cannot be conclusively proven. Serine is among the poorly conserved residues when mitochondrial amino acid sequences from different species are aligned. As previously reported in a comparison between human and bovine mitochondrial sequences⁸, tryptophan and arginine are among the most highly conserved amino acids. A convenient consequence is

Fig. 1 Sequence of the *D. melanogaster* mtDNA fragment. The main coding strand is shown containing the sense sequence of all mRNAs and most tRNAs encoded in the fragment; these are therefore transcribed from the complementary strand. The genes for cytochrome oxidase subunits I to III (CO I-III) and ATPase subunit 6, as well as the unassigned reading frames (URFs) 2 and A6L, are indicated; the 5' end of URF 2 and the 3' end of the CO III gene are outside the fragment. The genes are shown translated into the one-letter amino acid code with asterisks designating termination codons and cross-hatches indicating that no termination codon is present in the DNA sequence (see text). The tRNA genes are shown boxed. The sense of each gene is indicated by an arrow; -->, main coding strand. Base composition of the main coding strand: 41.5% T, 13.0% C, 33.7% A, 11.8% G. The sequence was derived from a 4,869-bp *Hind*III fragment of mtDNA from *D. melanogaster* (stock 1(1)E12¹⁸/C1B ♀) contained in the recombinant plasmid M2/8 (W. J. Gehring, unpublished results); the strain has an Oregon-R background (W. J. Gehring, personal communication). A 'shotgun' DNA sequencing strategy was used⁶² in which the mtDNA fragment was randomly fragmented by sonication⁶³ and subcloned in the vectors M13mp8 and M13mp9⁶⁴ to prepare single-stranded template. M13 recombinants were selected at random and the inserts sequenced by the dideoxynucleotide chain termination method^{65,66} using a flanking universal primer⁶⁷. A consensus sequence was assembled from individual insert sequences using the computer programs of Staden⁶⁸ to compile, edit and overlap the data. The sequence of the same mtDNA fragment derived from *D. melanogaster* (Oregon-R) and contained in recombinant plasmid pMD109²⁹, was determined for confirmatory purposes (see text). It contains 26 nucleotide changes compared with the sequence shown. These have been indicated in parentheses below that sequence in the corresponding positions. In those cases where such a nucleotide change resulted in an amino acid replacement, the new residue is shown in parentheses above the translated sequence.

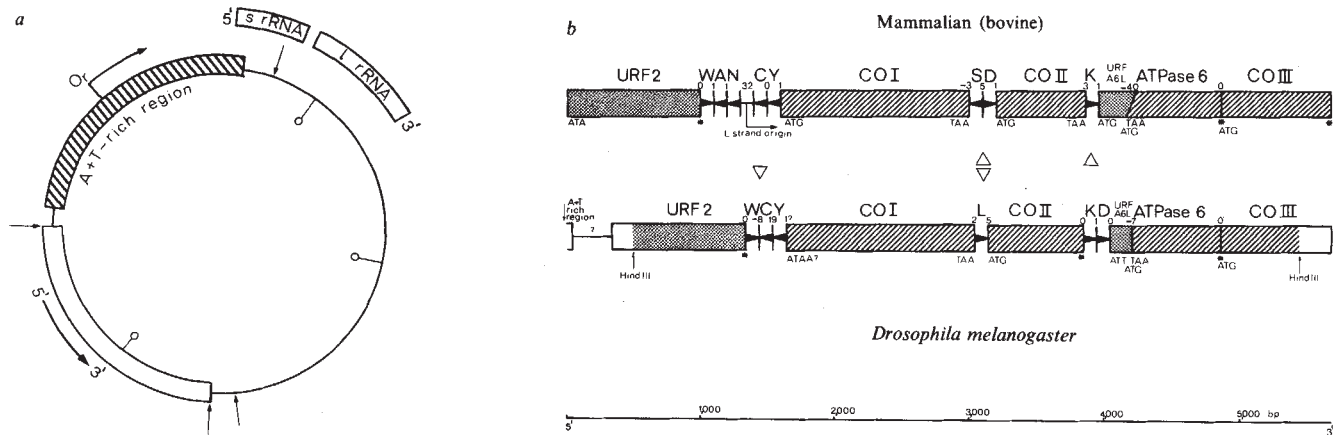


Fig. 2 Organization of the *D. melanogaster* mitochondrial genome. *a*, A circular map showing the position and orientation of the sequenced fragment (long open bar) relative to the *Hind*III (↑) and *Hae*III (○) recognition sites⁶. The positions of the A+T-rich region, origin of replication and rRNA genes, as well as the direction of DNA synthesis and transcription of rRNAs are as published^{2,6}. *b*, A gene map of the sequenced mtDNA fragment (Fig. 1) aligned with the homologous region of bovine mtDNA⁸. Protein-coding regions are represented by cross-hatched and URFs by shaded rectangles. The two *Hind*III recognition sites coincide with the ends of the fragment. The size of URF 2 and CO III have been estimated by alignment to the bovine genes. The tRNA genes are represented by > (main coding strand) or < and are identified by the one-letter amino acid code. Initiation and termination codons are shown and an asterisk indicates that no termination codon is present in the DNA sequence (see text). The numbers at the gene junctions denote the number of intergenic nucleotides at that point, with negative numbers signifying an overlap. Open triangles indicate regions of major difference between the two gene maps.

Fig. 3 Comparison of homologous protein-coding sequences in *D. melanogaster*, bovine⁸ and yeast²⁴ mtDNA. *a*, Alignment of the 5' ends of URF A6L. No sequence homologous to animal URF A6L has been reported for yeast. *b*, Alignment of the 5' ends of the gene coding for cytochrome oxidase subunit I. Asterisks between two sequences indicate the positions of identical nucleotides. The amino acid sequences have been deduced from the DNA sequence using the appropriate genetic codes^{7,28}. Sequences were aligned as described in the Table 1 legend. Initiation of translation for both *D. melanogaster* genes is discussed in the text.

a URF A6L	
Bovine	M P Q L D T S T W L T M I L S M F L T L F I I F Q
ATGCCGCACTAGACACGTCACATGACTGACATGATCTTATCAATATCTTGACCTTTTATCATCTTTCAA	
*** **	*** **
ATCCACAAATGACCTATTAGATGATTATTATTATTATTATT---TTTCTATTACATTTATTTATTTGT	
I P Q M A P I S W L L L F I I F S I T F I L F C	
<i>Drosophila</i>	
b CYTOCHROME OXIDASE, SUBUNIT I	
Bovine	M F I N R W L F S T N H K D I G T L Y L L F G A W
ATGTTTCATTACCGCTGACTATTCTCAACCAACCATAAAGATATTGGTACCCTTTATCTACTATTGGTGGCTTGG	
*** **	*** **
ATAA--TCGCGACAATGATTATTTCTACAAATCATAAAGATATTGGAACTTTATATTATTTTGGAGCTTGA	
M? S R Q W L F S T N H K D I G T L Y F I F G A W	
<i>Drosophila</i>	
Yeast	M U Q R W L Y S T N A K D I A U L Y F M L A I F
ATG-GTACAAGATGATTATTTCAACAAATGCAAAAGATATTGCAGTATTATTTTATGTTAGCTATTTT	
*** **	*** **
ATAATCGCGCAATGATTATTTCTACAAATCATAAAGATATTGGAACTTTATATTATTTTGGAGCTTGA	
M? S R Q W L F S T N H K D I G T L Y F I F G A W	
<i>Drosophila</i>	

URF A6L leaves no doubt that the former starts with the isoleucine codon ATT immediately distal to the tRNA^{Asp} gene (Fig. 3a). There are precedents for this in both human and mouse mitochondria^{7,9}.

Comparison of the *Drosophila*, bovine and yeast CO I genes (Fig. 3b) shows that the ATG initiation codons in the bovine and yeast genes align with an in-phase TAA termination codon in the *Drosophila* gene. Two trivial reasons for this, a sequencing error or a cloning artefact, can be excluded. The region has been sequenced several times on both strands. Also, analysis of the corresponding mtDNA fragment from a different *D. melanogaster* strain, inserted in a different vector²⁹, shows it to have an identical sequence (Fig. 1). The reading frame could represent a pseudogene with an active CO I gene existing elsewhere, but this is unlikely. As shown in Fig. 1, all nucleotide changes between the CO I genes of the two *D. melanogaster* strains represent silent substitutions, in contrast to those in URF 2 and the ATPase 6 gene, some of which are replacement substitutions. As no appreciable sequence drift has occurred between these two genes, inactivation would have had to be a recent event (in evolutionary terms), but this is excluded because the genes in both strains are affected.

The high homology of the putative gene to its mammalian and fungal counterparts (see Table 1) makes it unlikely that translation is initiated at any codon downstream from the TGA

codon specifying the first highly conserved amino acid (tryptophan). Homology may be extended upstream of the tryptophan codon because the amino acid sequence V-Q-R in yeast CO I is very similar to S-R-Q in *Drosophila* (Fig. 3b). The possibility that the serine codon TCG acts as an initiator cannot be discounted even though it is used once as an internal codon in URF 2 and aligns with serine in bovine URF 2. Curiously, also *D. melanogaster* CO II and III, but none of the previously published mitochondrially encoded cytochrome oxidase sequences^{7-9,23-25,30-33}, have a serine residue immediately after formyl-methionine but specified by a different codon. As the preceding cysteine and tyrosine tRNA genes are on the complementary strand, a spliced CO I gene could start as many as 145 nucleotides upstream from TCG. This would be at variance with the otherwise simple pattern of transcription and post-transcriptional processing, which *Drosophila* mitochondria seem to have in common with their mammalian counterparts (see below). Also, the region contains no ATG codons and no reading frames displaying homology to any of the known N-terminal CO I sequences. It therefore appears that initiation of translation would have to take place at or very near to the anomalous TAA termination codon.

An ATA codon, almost certainly an initiation codon in mammalian mitochondria^{7,17}, overlaps the TAA codon in the -1 frame. Here, this frame does not contain an open reading frame

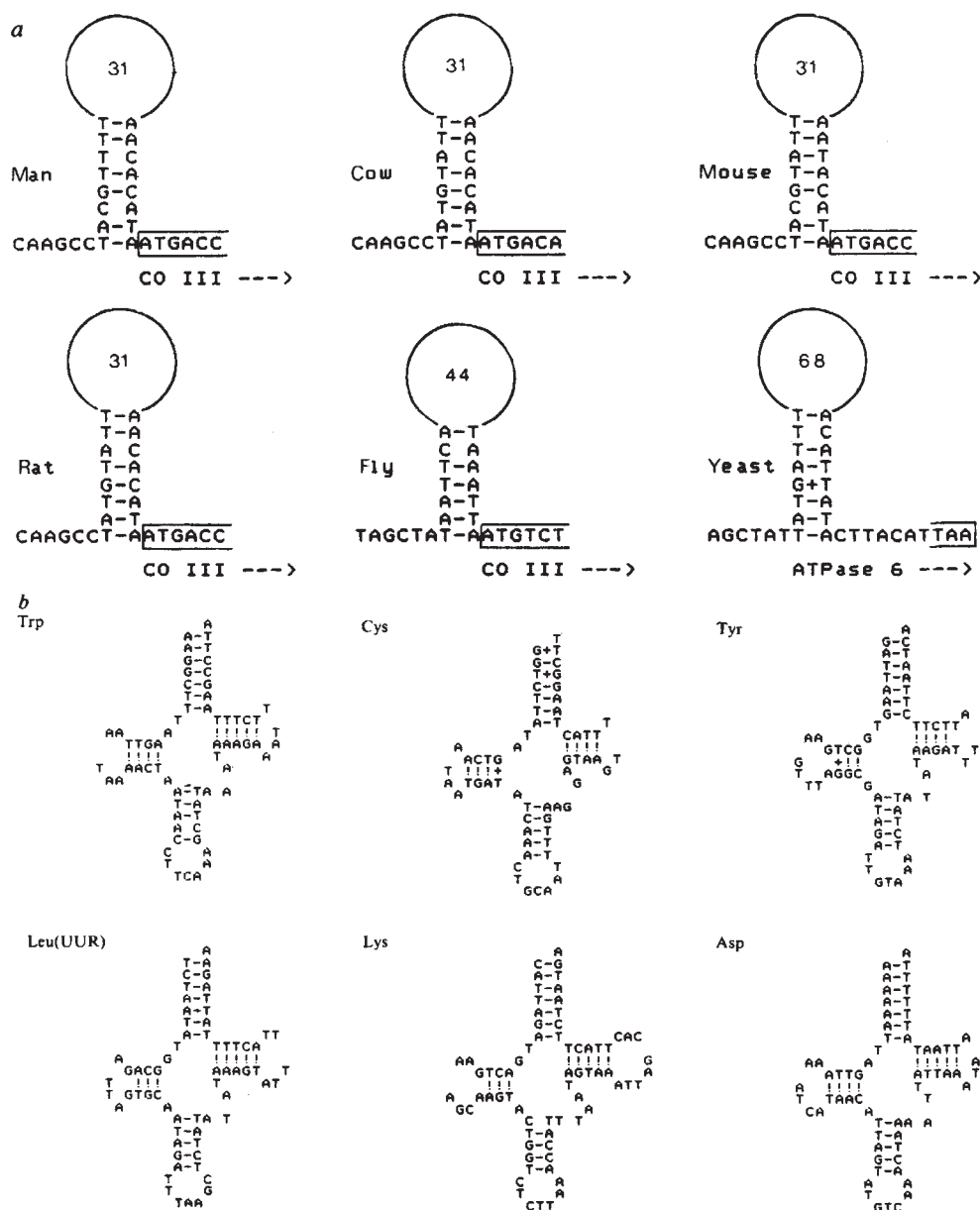


Fig. 4 Predicted RNA secondary structures. *a*, Potential secondary structures encoded at the 3' end of the ATPase 6 genes of man⁷, cow⁸, mouse⁹, rat³⁰, *D. melanogaster* and the yeast *Saccharomyces cerevisiae*²⁶. The DNA sequences have been drawn as a loop-and-stem structure. The loop sequences are represented by the loop on top of each stem and their size, in nucleotides, is indicated. In all animals, but not in yeast, the 3' end of the stem is adjacent to the 5' end of the CO III gene. The significance of this is discussed in the text. *b*, Secondary structures of the six tRNAs encoded in the *D. melanogaster* mtDNA fragment. The DNA sequences of the six genes have been drawn in the form of a cloverleaf. At the upper left of each tRNA the amino acid is indicated which it is thought to accept. Since there are two codon families for leucine (Fig. 1), the codons read by the leucyl-tRNA are specified. The symbols - and + indicate Watson-Crick base pairs and + denotes a G-U pair.

but ATA could function in the highly conserved 0 frame, if it were part of a 4-nucleotide initiation codon ATAA or if initiation were followed by a 1-nucleotide ribosomal frameshift. Site-specific ribosomal frameshifting during translation has been observed in coliphage f2 between the frames of the coat and lysis proteins³⁴. Frameshift mutants with a leaky phenotype due to corrective ribosomal frameshifting have been widely reported³⁵⁻⁴¹ and in one case involve mutant CO II genes in yeast mitochondria⁴². Little is known about how ribosomal frameshifts or frameshift suppressions (corrective ribosomal frameshifts) occur, but in many cases tRNAs seem to be implicated. For example, an induced frameshift mutation in the *hisD* gene of *Salmonella typhimurium*, causing a change from a glycine codon GGG to GGGG, was suppressed by a corresponding change in the glycine tRNA anticodon from CCC to CCCC^{37,43}. This could be explained by quadruplet reading or by triplet reading with the extra base in the anticodon loop acting as a spacer, causing steric hindrance to the tRNA reading the following triplet^{44,45}. Frameshifts could be due to mutations affecting not the anticodon but another part of the tRNA. The distance over which mRNA moves in each decoding step is about 10 Å, which is small compared with the dimensions of a tRNA⁴¹. A small change in tRNA structure could, therefore, cause significant strain on the interaction of a second tRNA with the neighbouring codon. Even wild-type tRNAs have been

implicated in the corrective frameshifting of frameshift mutants with a leaky phenotype⁴⁶.

A putative tRNA^{Met} in *D. yakuba* mitochondria has been reported¹³. If conserved in *D. melanogaster*, it would be able to recognize the codon ATA (by analogy to the situation in mammalian mitochondria⁷), but not ATAA. Further sequencing of the *D. melanogaster* mitochondrial genome, now in progress, might provide for a highly unusual initiator tRNA having a quadruplet anticodon or an anomalous structure. It is interesting that one of the yeast mitochondrial threonine tRNAs has an 8 rather than 7-nucleotide anticodon loop⁴⁷. The fact that it still recognizes triplet codons could be related to its specific pattern of base modifications⁴⁸.

Transcription/post-transcriptional processing

In the transcript map of *D. melanogaster* mtDNA, most polyadenylated transcripts and the rRNAs map adjacent to each other without overlapping¹². Size and hybridization behaviour of some larger transcripts are consistent with a precursor function. They are homologous to the sequenced mtDNA fragment, with a common end near the 5' end of that fragment. The available data are consistent with the protein-coding strand being transcribed completely as a large primary transcript which is processed according to the 'tRNA punctuation model' proposed for human mitochondria⁴⁹.

Table 2 Amino acid assignments for the codons AGA and ATA in *D. melanogaster* mitochondria

Codon	Amino acid	B	Y	B/Y
AGA	Serine	9	6	3
	Threonine	6	2	0
	Alanine	2	5	0
	Other	6	5	1
ATA	Methionine	39	16	14
	Leucine	11	7	2
	Isoleucine	5	8	2
	Threonine	8	3	0
	Other	19	17	0

The numbers denote the frequency with which the codons AGA and ATA in the *D. melanogaster* mtDNA fragment align with codons assigned to specified or other amino acids in the corresponding bovine (B) or yeast (Y) sequence, or in both (B/Y). The compilation was made using the sequence alignments described in Table 1 legend.

The transcript map is partly consistent with the gene organization derived from the DNA sequence. Size and hybridization behaviour leave no doubt that transcripts B3 and B11 are homologous to the CO I and CO II genes, respectively, and that the transcript doublet B8/9 corresponds to the ATPase 6 and CO III genes. Merten and Pardue¹² report a low-abundance transcript (B6) for which mapping is ambiguous. B6 hybridized to both the CO I and large rRNA coding regions and they concluded that a sequence homologous to B6 is repeated in the genome. An alternative interpretation, however, is that B6 represents a low abundance doublet. The size of B6 is consistent with transcripts from both URF 1 and URF 2 and a B6 doublet would be consistent with the gene order reported for *D. yakuba* mtDNA¹³.

Termination of translation

Only the CO I gene has a termination codon (TAA) that does not overlap with a neighbouring gene on the same strand. If the tRNA punctuation model of RNA processing is operative, the URF 2 and CO II transcripts would end with a solitary U following the C-terminal amino acid codon and the ATPase 6 transcript with UA (Fig. 1). Because mitochondrial transcripts in *D. melanogaster* are polyadenylated¹¹, in-phase UAA termination codons would be created post-transcriptionally in all three transcripts. The polyadenylation model for termination of translation⁵⁰ which has been shown to function in mammalian mitochondria⁵¹ thus appears to operate in *Drosophila* mitochondria as well.

There are two instances where this model does not fit the experimental data. The first is the junction between URF A6L and the overlapping ATPase 6 gene, where no specific signal for endonucleolytic cleavage is apparent. Both the *D. melanogaster* and mammalian mitochondrial transcript maps^{12,52} suggest that URF A6L and the ATPase 6 gene are transcribed as a unit and possible ways in which such a multicistronic message may be translated have been discussed previously⁸.

A second anomaly concerns the junction between the ATPase 6 and CO III genes. As in other animal mitochondrial genomes⁷⁻⁹, the TAA termination codon of ATPase 6 overlaps with the ATG initiation codon of CO III by one nucleotide. Transcript maps^{12,51-53} leave no doubt that two individual transcripts are created by cleavage of a primary transcript at the A preceding the AUG initiation codon of CO III followed by polyadenylation which restores the UAA termination codon of ATPase 6. However, no tRNA gene is detectable which could provide a recognition signal for such cleavage. Since the general cleavage pattern observed also implicates antisense tRNA cloverleaf structures in recognition^{7,51,53}, an exact tRNA cloverleaf might not be required. Bibb *et al.*⁹ point out that the mouse ATPase 6 sequence is capable of forming a loop-and-stem

structure immediately proximal to the initiation codon of CO III. Equivalent secondary structures are possible in both other mammalian sequences and that of *D. melanogaster* (Fig. 4a), but not in those from yeast and *Aspergillus* where the 'tRNA punctuation' model does not apply^{26,32,54}.

tRNA genes

The identified *D. melanogaster* tRNAs (Fig. 4b and ref. 13) closely resemble their mammalian counterparts in lacking the universal sequence T-ψ-C-R-A, the constant 7-nucleotide size of the 'TψC' loop, and a number of universal bases in the D-loop and elsewhere^{7,8,55}. The universal 3'-terminal C-C-A sequence is not encoded. A striking feature of mitochondrial tRNAs appears to be their high A+U content⁵⁶. The *D. melanogaster* mitochondrial tRNA^{Asp} with an A+U content of 90% is the most extreme reported example of this.

Whereas tRNA^{Leu}_{UUR} in mammalian mitochondria is unique in having most of the invariant and semi-invariant features of non-mitochondrial tRNAs, the *D. melanogaster* variant does not. As the mammalian tRNA^{Leu}_{UUR} genes are located at a position in the genome that has been proposed to contain a site for termination or attenuation of transcription⁵⁷ and the *D. melanogaster* variant is located elsewhere, it is tempting to suggest that retention of universal tRNA features by the mammalian species is in some way related to a specific additional function of this tRNA or its gene.

The tRNA^{Lys} differs from its mammalian mitochondrial counterparts in having the anticodon CUU rather than UUU. This appears to restrict codon recognition to AAG instead of AAR. A second lysine tRNA capable of decoding AAA might be encoded or imported. There are precedents for both in fungal mitochondria^{58,59}. Alternatively, tRNA^{Lys} could be capable of decoding both AAA and AAG, by analogy to recognition of AUA and AUG by the mammalian mitochondrial initiator tRNA with anticodon sequence CAU⁷.

I thank Drs W. J. Gehring and J. O. Bishop for providing the recombinant plasmids M2/8 and pMD109, respectively, and Drs B. G. Barrell and F. Sanger for critically reading the manuscript.

Received 25 March; accepted 27 May 1983.

1. Fauron, C. M.-R. & Wolstenholme, D. R. *Proc. natn. Acad. Sci. U.S.A.* **73**, 3623-3627 (1976).
2. Peacock, W. J. *et al. Cold Spring Harb. Symp. quant. Biol.* **38**, 405-416 (1973).
3. Shah, D. M. & Langley, C. H. *Nature* **281**, 696-699 (1979).
4. Fauron, C. M.-R. & Wolstenholme, D. R. *Nucleic Acids Res.* **8**, 2439-2452 (1980).
5. Goddard, J. M. & Wolstenholme, D. R. *Nucleic Acids Res.* **8**, 741-757 (1980).
6. Klukas, C. K. & Dawid, I. B. *Cell* **9**, 615-625 (1976).
7. Anderson, S. *et al. Nature* **290**, 457-465 (1981).
8. Anderson, S. *et al. J. molec. Biol.* **156**, 683-717 (1982).
9. Bibb, M. J., Van Etten, R. A., Wright, C. T., Walberg, M. W. & Clayton, D. A. *Cell* **26**, 167-180 (1981).
10. Ramirez, J. L. & Dawid, I. B. *J. molec. Biol.* **119**, 133-146 (1978).
11. Spradling, A., Pardue, M. L. & Penman, S. *J. molec. Biol.* **109**, 559-587 (1977).
12. Merten, S. H. & Pardue, M. L. *J. molec. Biol.* **153**, 1-21 (1981).
13. Clary, D. O., Goddard, J. M., Martin, S. C., Fauron, C. M.-R. & Wolstenholme, D. R. *Nucleic Acids Res.* **10**, 6619-6637 (1982).
14. Staden, R. *Nucleic Acids Res.* **8**, 817-825 (1980).
15. Barrell, B. G., Bankier, A. T. & Drouin, J. *Nature* **282**, 189-194 (1979).
16. Martens, P. A. & Clayton, D. A. *J. molec. Biol.* **135**, 327-351 (1979).
17. Young, I. G. & Anderson, S. *Gene* **12**, 257-265 (1980).
18. Macino, G., Coruzzi, G., Nobrega, F. G., Li, M. & Tzagoloff, A. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3784-3785 (1979).
19. Yin, S. *et al. in Mitochondrial Genes* (eds Slonimski, P., Borst, P. & Attardi, G.) 361-373 (Cold Spring Harbor Laboratory, New York, 1982).
20. Waring, R. B. *et al. Cell* **27**, 4-11 (1981).
21. Bonitz, S. G. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 3167-3170 (1980).
22. Heckman, J. E., Sarnoff, J., Alzner-DeWeerd, B., Yin, S. & RajBhandary, U. L. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3159-3163 (1980).
23. Fox, T. D. *Proc. natn. Acad. Sci. U.S.A.* **76**, 6534-6538 (1979).
24. Bonitz, S. G., Coruzzi, G., Thalenfeld, B. E., Tzagoloff, A. & Macino, G. *J. biol. Chem.* **255**, 11927-11941 (1980).
25. Thalenfeld, B. E. & Tzagoloff, A. *J. biol. Chem.* **255**, 6173-6180 (1980).
26. Macino, G. & Tzagoloff, A. *Cell* **20**, 507-517 (1980).
27. Dickerson, R. E. & Geis, I. *The Structure and Action of Proteins*, 19-20 (Benjamin/Cummings, Menlo Park, 1969).
28. Hudspeeth, M. E. S., Ainley, M. W., Shumard, D. S., Butow, R. A. & Grossman, L. I. *Cell* **30**, 617-626 (1982).
29. Mason, P. J. & Bishop, J. O. *Biochem. biophys. Res. Commun.* **95**, 1268-1274 (1980).
30. Grosskopf, R. & Feldmann, H. *Curr. Genet.* **4**, 151-158 (1981).
31. Browning, K. S. & RajBhandary, U. L. *J. biol. Chem.* **257**, 5253-5256 (1982).
32. Netzer, R., Köchel, H. G., Basak, N. & Kuntzel, H. *Nucleic Acids Res.* **10**, 4783-4794 (1982).
33. Fox, T. D. & Leaver, C. J. *Cell* **26**, 315-323 (1981).

34. Beremand, M. N. & Blumenthal, T. *Cell* **18**, 257–266 (1979).
35. Donahue, T. F., Farabaugh, P. J. & Fink, G. R. *Science* **212**, 455–457 (1981).
36. Riddle, D. L. & Roth, J. R. *J. molec. Biol.* **66**, 495–506 (1972).
37. Riddle, D. L. & Carbon, J. *Nature new Biol.* **242**, 230–234 (1973).
38. Atkins, J. F. in *Transfer RNA: Biological Aspects* (eds Söll, D., Abelson, J. N. & Schimmel, P. R.) 439–449 (Cold Spring Harbor Laboratory, New York, 1979).
39. Atkins, J. F. & Ryce, S. *Nature* **249**, 527–529 (1974).
40. Bossi, L. & Roth, J. R. *Nature* **286**, 123–127 (1980).
41. Kurland, C. G. in *Nonsense Mutations and tRNA Suppressors* (eds Celis, J. E. & Smith, J. D.) 97–108 (Academic, New York, 1979).
42. Fox, T. D. & Weiss-Brummer, B. *Nature* **288**, 60–63 (1980).
43. Yourno, J. *Nature new Biol.* **239**, 219–221 (1972).
44. Bossi, L. & Roth, J. R. *Cell* **25**, 489–496 (1981).
45. Gaber, R. F. & Culbertson, M. R. *Gene* **19**, 163–172 (1982).
46. Atkins, J. F., Gesteland, R. F., Reid, B. R. & Anderson, C. W. *Cell* **18**, 1119–1131 (1979).
47. Macino, G. & Tzagoloff, A. *Proc. natn. Acad. Sci. U.S.A.* **76**, 131–135 (1979).
48. Sibley, A.-P., Dirheimer, G. & Martin, R. P. *FEBS Lett.* **132**, 344–348 (1981).
49. Attardi, G. et al. in *ICN-UCLA Symp. Extrachromosomal DNA* (eds Cummings, D., Borst, P., Dawid, I., Weissman, S. & Fox, C. F.) 443–469 (Academic, New York, 1979).
50. Barrell, B. G. et al. in *31st Mosbach Colloq. Biological Chemistry of Organelle Formation* (eds Bücher, T. H., Sebald, W. & Weiss, H.) 11–25 (Springer, Berlin, 1980).
51. Ojala, D., Montoya, J. & Attardi, G. *Nature* **290**, 470–474 (1981).
52. Ojala, D., Merkel, C., Gelfand, R. & Attardi, G. *Cell* **22**, 393–403 (1980).
53. Montoya, J., Ojala, D. & Attardi, G. *Nature* **290**, 465–470 (1981).
54. Borst, P. & Grivell, L. A. *Nature* **290**, 443–444 (1981).
55. Dirheimer, G., Keith, G., Sibley, A.-P. & Martin, R. P. in *Transfer RNA: Structure, Properties and Recognition* (eds Schimmel, P. R., Söll, D. & Abelson, J. N.) 19–41 (Cold Spring Harbor Laboratory, New York, 1979).
56. Eigen, M. & Winkler-Oswatitsch, R. *Naturwissenschaften* **68**, 217–228 (1981).
57. Dubin, D. T., Montoya, J., Timko, K. D. & Attardi, G. *J. molec. Biol.* **157**, 1–19 (1982).
58. Köchel, H. G., Lazarus, C. M., Basak, N. & Kuntzel, H. *Cell* **23**, 625–633 (1981).
59. Martin, R. P., Schneller, J.-M., Stahl, A. J. C. & Dirheimer, G. *Biochemistry* **18**, 4600–4605 (1979).
60. Dayhoff, M. O. *Atlas of Protein Sequence and Structure* (National Biomedical Research Foundation, Silver Springs, Maryland, 1969).
61. Staden, R. *Nucleic Acids Res.* **4**, 4037–4051 (1977).
62. Sanger, F., Coulson, A. R., Barrell, B. G., Smith, A. J. H. & Roe, B. A. *J. molec. Biol.* **143**, 161–178 (1980).
63. Deininger, P. *Analyt. Biochem.* **129**, 216–223 (1983).
64. Messing, J. & Vieira, J. *Gene* **19**, 269–276 (1982).
65. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463–5467 (1977).
66. Sanger, F. & Coulson, A. R., *FEBS Lett.* **87**, 107–110 (1978).
67. Duckworth, M. L. et al. *Nucleic Acids Res.* **9**, 1691–1706 (1981).
68. Staden, R. *Nucleic Acids Res.* **10**, 4731–4751 (1982).

LETTERS TO NATURE

Rapid variations in forbidden emission line profiles in a Seyfert 2 galaxy

A. P. Fairall

Department of Astronomy, University of Cape Town,
Rondebosch 7700, South Africa

F-427 (ref. 1) = ESO263 – G13 (ref. 2) is seen as a nearly flat-on spiral galaxy about 55 arc s in diameter on the SERC IIIa-J sky survey. It has a bright and distinct nucleus and has been selected for a survey of compact and bright-nucleus galaxies. Spectroscopic observations at the South African Astronomical Observatory, Sutherland of this extreme Seyfert 2 galaxy (RA 10 h 07 min 44 s, dec $-42^{\circ} 33' 30''$) reveal apparent profile variations in the [O III] 5,007 + 4,959 emission lines over as short an interval as 30 min—in complete contradiction to theoretical considerations. Attempts to explain the variations as being statistical, instrumental or spatial in origin are unsuccessful.

All the observations were made using the 1.9-m Radcliffe reflector with UNIT spectrograph and Reticon Photon Counting System³ at Sutherland. Initially, a low-dispersion (200 Å mm^{-1}) spectrogram and high-dispersion (50 Å mm^{-1}) spectrograms of the [O III] 5,007 region and the H α region were obtained in February 1982. In the low-dispersion spectrogram (Fig. 1), extremely strong [O III] 5,007 + 4,959 lines, together with [O II] 3,727, [Ne III] 3,870, [N II] 6,584 + 6,548, the Balmer lines H α to H γ and He II 4,686, are present. The [O III] lines are broadened with FWZI = 900 km s^{-1} (FWHI = 360 km s^{-1}) thus making this a Seyfert 2 galaxy^{4–6}. Seyfert 2 galaxies are also characterized by a high 5,007/H β (or 4,959/H β) ratio^{7–9}. In the present galaxy 5,007/H β = 19, an indication of unusually high excitation which is supported by the strength of the He II 4,686 line (4,686/H β = 0.6). The redshift of the galaxy corresponds to $V = 9,980 \text{ km s}^{-1}$, hence $V_0 = V + 300 \sin i \cos b = 9,630 \text{ km s}^{-1}$.

Due to the weakness of the higher lines, the Balmer decrement cannot be measured accurately, but H α /H β = 5.3 which indicates (using standard values and case B) an extinction A_V of 1.9 mag. The slope of the continuum is $\alpha \sim 3.4$ (where $F_\nu \propto \lambda^\alpha$) which also suggests reddening. If the intrinsic slope were $\alpha = 0.7$ (typical synchrotron) the extinction $A_V = 2.2$ mag, or higher if the intrinsic slope included thermal contributions. The continuum extinction may be higher than the Balmer extinction, if the former originates deeper inside a dusty nuclear region.

The high-dispersion spectra (obtained in 1982) showed the [O III] lines and the [N II] 6,584 line to have a double peak—an unusual, if not unique, profile in a Seyfert forbidden emission line. To confirm this structure, further high-dispersion

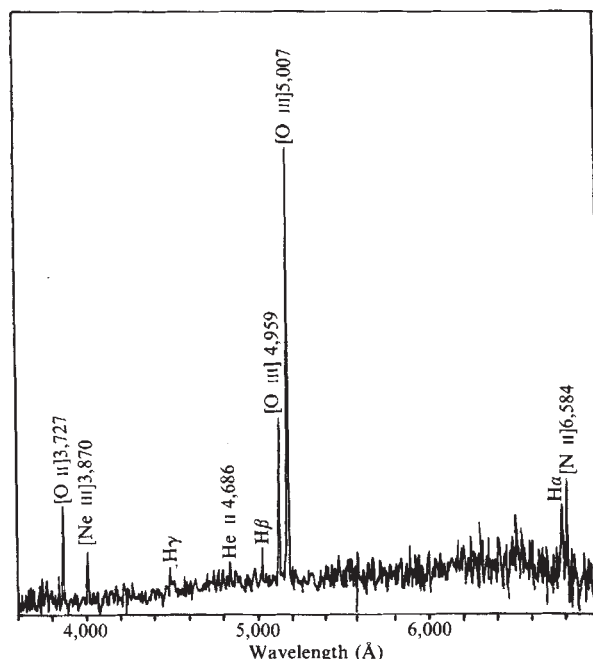


Fig. 1 The low-dispersion spectrum. The vertical scaling has been corrected for instrumental response (but there is some unevenness at the red end due to a low number of calibration points).

Table 1 Observing log

	Exposure UT (mid exp)	N
	1983	
Night 1	18–19 January A	00 h 21 min 280
	B	00 49 70
	C	01 19 60
	D	01 48 70
	E	02 16 280
Night 2	19–20 January F	21 54 460
Night 3	20–21 January G	21 44 400
	H	22 11 400
Night 4	21–22 January J	01 01 250
Night 5	22–23 January K	23 47 110
Night 6	23–24 January L	22 32 280
	M	22 59 370
	N	23 27 270
	P	22 39 330
	Q	23 08 320
	R	23 36 270
	S	00 04 270
	T	00 32 280
Night 7	24–25 January U	01 00 240

All exposures 25 min. N = height of 5,007 profile (the highest number of counts per pixel) above continuum level.

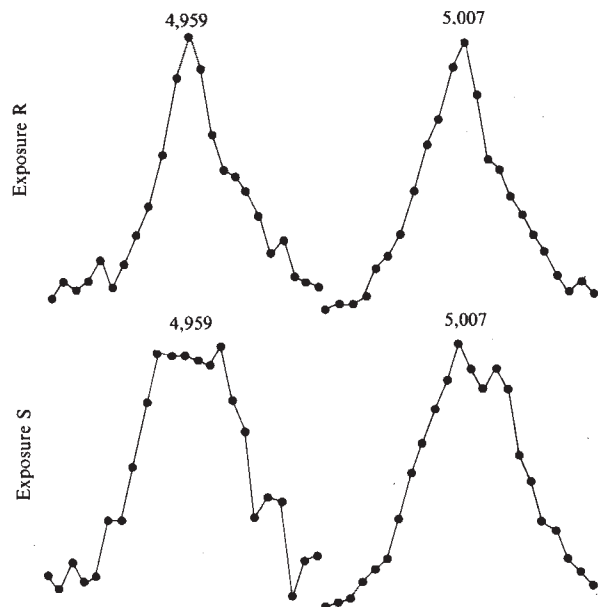


Fig. 2 Selected line profiles from Night 7 that show a conspicuous variation in profile from one exposure to the next. The shape of the 5,007 profile is confirmed by the noisier 4,959 line from the same exposure. Statistical comparison is discussed in the text. Pixel spacing is $0.6 \text{ \AA}$ (35 km s^{-1}), the 4,959 profiles have been scaled vertically to match the height of the 5,007 lines. The width of the 5,577 sky line on both exposures is 3 pixels.

spectrogram exposures were obtained in January 1983. Suspicions of line profile variability led to 19 exposures being obtained over seven consecutive nights (see Table 1; although all nights were clear, there is considerable fluctuation in the seeing).

Variations in profile widths and structure are apparent from night to night, and even over consecutive exposures. Figure 2 shows some of the best evidence. What is apparent to the eye is confirmed by a χ^2 test:

Using the central 18 pixels (that is, 17 degrees of freedom)

Exposure R 5,007/4,959	$\chi^2 = 31$
Exposure S 5,007/4,959	$\chi^2 = 20$
Exposure R 5,007/Exposure S 5,007	$\chi^2 = 53$

The first two results are acceptable as being derived from the same profile (although the first gives a higher χ^2 value than those derived from other 5,007/4,959 samples). However, in the third result, the probability of the profile from exposure R being the same as that from exposure S is virtually zero ($\chi^2 = 36$ gives a probability of 0.005). Thus, from a statistical point of view, there seems little doubt that the variation is real.

Figure 3 shows 5,007 profiles from all the exposures. The four other exposures from night 7 do not show significant variations (χ^2 values do not exceed 30). Night 1, although plagued by poor seeing, is similar to night 7 (each has one exposure showing the broader profile). Several exposures from the other nights show the broader profile, usually with the double peak. The night 2 profile matches that found in 1982 (not shown). While it seems that it is the redward peak that grows and collapses, the asymmetry of profiles from night 3, and one from night 6, suggests that the blueward peak may also vary. The broad base of the line remains stable and the variations affect at most some 35% of the total line flux.

Could the variations be instrumental in origin? On exposures G to U, the 5,577 night sky line is available; its profile is found to be consistent (3 pixels wide, the slit width being set at 1.8 arc s). An instrumental effect could hardly manipulate the galaxy lines without also affecting the sky lines. Also, all intervening calibration arcs have been examined and all arc line widths are consistent. All exposures are corrected for individual response by a 'flat field', and the same flat field is used for all exposures from the same night (yet variations occur). All the

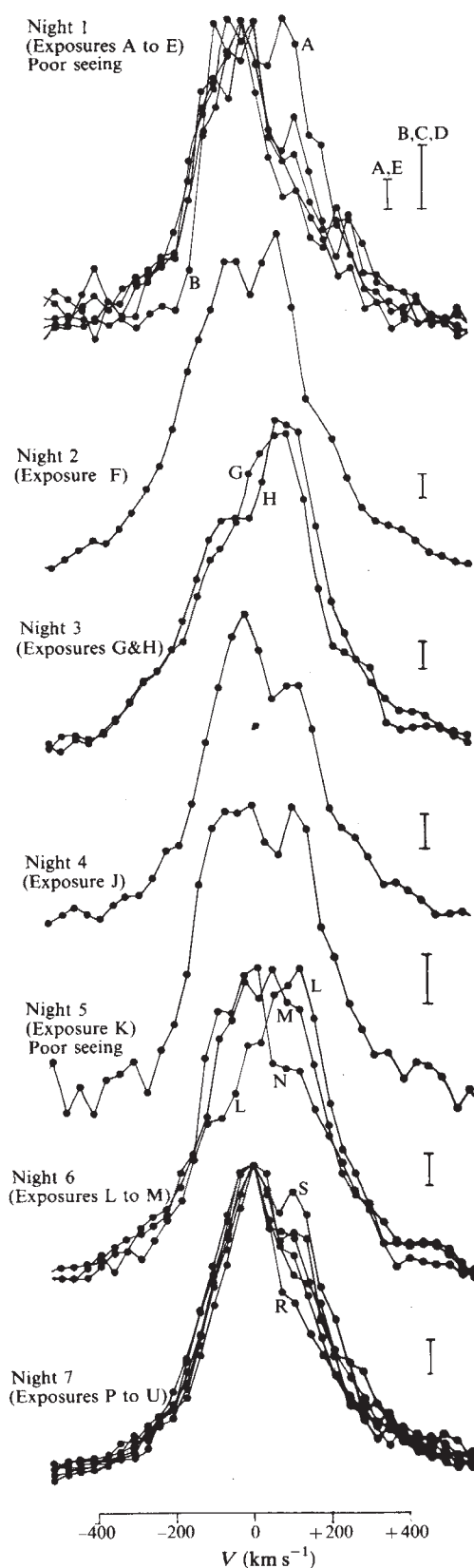


Fig. 3 Profiles of the 5,007 line from all 19 exposures made in January 1983. The heights of the lines have been normalized to compensate for seeing variations; best results came from matching the height of the blueward peak (except for exposures G, H and L). Even if normalization factors are questioned, the profile shapes (confirmed by the 4,959 lines from the same exposures) are seen to vary. Photon counts in the highest pixels are given (as N) in Table 1. The error bars are 2σ where $\sigma = \sqrt{2/3 N}$, and are shown as a rough indication of expected noise.

variations shown in Figs 2 and 3 can be seen in the flat field, but otherwise unreduced, raw data, so that they are not a product of the reduction procedures. Also, due to grating changes, the lines are never recorded in exactly the same pixels from night to night: on certain nights a totally different portion of the Reticon array was used. A second Reticon array records the sky 30 arc s from the galaxy nucleus. All sky spectra have been examined and appear normal with no trace of contamination.

The spatial variations still need to be excluded. The image of the galaxy was held steady on the slit by offset guiding using a convenient star only 15 arc s west-southwest of the galaxy nucleus. Exactly the same offset was used for all exposures.

The galaxy has a distinct nucleus. Examination of its image on the SERC IIIa-J and, more especially, on the ESO Quick Blue Survey, using high magnification does not reveal any multiplicity or structure to the nucleus. At the telescope it was observed using the integration television system and showed up as a single image, indistinguishable from a star. Increasing the gain and integration well above normal setting failed to reveal any surrounding disk nor did it reveal any structure to the nucleus. Seeing would further blur the image, so there would be no chance of spatial resolution. Variations in the profile show up during very poor seeing (Tuesday) just as they show up during good seeing (Monday).

During the same observing run some multiple exposures were made of the emission lines of another (Seyfert 1) galaxy. More extensive multiple exposures of that galaxy were made 4 months earlier. No profile variations, similar to those reported here, were found in those data.

Thus, in the absence of any alternative cause, it would seem that the changes in the line profiles of F-427 = ESO263-G13 are real.

Variations in Balmer line profiles over longer time spans are well known in Seyfert galaxies. Long-term variation is more the rule than the exception and numerous examples are found in the literature¹⁰⁻¹⁷. However, in almost all cases, the periods concerned are of the order of years, the shortest mentioned being from month to month. Very few Seyfert galaxies show any structure in their forbidden lines; apparent changes in the strengths of forbidden lines are normally ascribed to continuum variations instead, although I have found an exception¹⁸.

Two major problems arise in the interpretation of the rapid variability. First, one would expect that [O III] emission cannot change faster than the time scale for changes in the photo-ionization rate to be transmitted, that is, the H recombination time. This is $\sim 10^{12}/n_e$ s, so even for $n_e = 10^6$ one would expect [O III] changes to be smeared over ~ 10 days. The second concerns the implied dimensions of the emitting region. A crude estimate of the strength of the 5,007 line gives 8×10^{-13} erg s⁻¹ cm⁻¹. If $H = 50$ km s⁻¹ Mpc⁻¹ and one restricts the argument to the 35% of the line responsible for the variation, this requires some 10^{42} erg s⁻¹ to be emitted from the region responsible (that is, 10^{53} photons s⁻¹). Assuming normal oxygen abundances, then even for $n_e \sim 10^6$ cm⁻³ the volume of emitting gas is $\sim 10^{56}$ cm³. From light travel time arguments, fast variations should not occur unless the geometry is extremely peculiar. As the quenching density is $\sim 10^4$ cm⁻³, we cannot reasonably appeal to higher densities. Thus it would seem that no conventional physical conditions can explain the observations. While this presents an enigma, it is not the first in the study of active nuclei.

Six further spectrograms of the 5,007 line have been obtained (on six successive nights) in May 1983 by Meaburn and Disney using the same equipment. The object was at times deliberately shifted around in the slit as a precaution against spatial resolution. Preliminary results confirm the profile variations. On two of the spectra, the redward peak is very high; there is also an additional variable blueward peak (blueshifted ~ 300 km s⁻¹).

I thank the director of the South African Astronomical Observatory, Professor M. W. Feast, for observing time and use of computer reduction facilities. L. Lowe and F. van Wyk

assisted at the telescope. I particularly thank Professor M. Disney as well as Dr J. Menzies and Professors J. M. Juritz and E. Capriotti for their efforts in checking my data and the reduction procedures. I also thank Dr J. Meaburn for allowing the confirmatory observations. I have had useful discussions with Drs A. Walker and T. Lloyd-Evans.

Received 17 March; accepted 27 May 1983.

1. Fairall, A. P. *Mon. Not. R. astr. Soc.* **203**, 47-54 (1983).
2. Holmberg, E. B., Lauberts, A., Schuster, H.-E. & West, R. M. *Astr. Astrophys. Suppl.* **22**, 295-342 (1977).
3. Jordan, A. R., Read, P. D. & van Breda, I. *Pap. at SPIE Conf. Instrumentation in Astronomy*, Tucson (1982).
4. Khachikian, E. Ye. & Weedman, D. W. *Astrophys. J.* **192**, 581-589 (1974).
5. Weedman, D. W. *A. Rev. Astr. Astrophys.* **15**, 69-95 (1977).
6. Koshi, A. T. *Astrophys. J.* **233**, 56-73 (1978).
7. Bohuski, T. J., Fairall, A. P. & Weedman, D. W. *Astrophys. J.* **221**, 776-787 (1978).
8. Fairall, A. P. *Mon. Not. R. astr. Soc.* **191**, 391-397 (1980).
9. Shuder, J. M. & Osterbrock, D. E. *Astrophys. J.* **250**, 55-65 (1981).
10. Tohline, J. E. & Osterbrock, D. E. *Astrophys. J. Lett.* **210**, L117 (1976).
11. Osterbrock, D. E., Koshi, A. T. & Phillips, M. M. *Astrophys. J.* **206**, 898-909 (1976).
12. Collin-Souffrin, S., Alloin, D. & Andrillat, Y. *Astr. Astrophys.* **22**, 343-354 (1973).
13. Boksenberg, A. & Netzer, H. *Astrophys. J.* **212**, 37-51 (1977).
14. Phillips, M. M. *Astrophys. J. Suppl.* **38**, 187-203 (1978).
15. Schulz, H. & Rafanelli, P. *Astr. Astrophys.* **103**, 216-219 (1981).
16. Foltz, C. B. *et al.* *Astrophys. J.* **250**, 508-512 (1981).
17. Peterson, B. M., Foltz, C. B., Byard, P. L. & Wagner, R. M. *Astrophys. J. Suppl.* **49**, 469-486 (1982).
18. Fairall, A. P. *Mon. Not. R. astr. Soc. S.Afr.* **39**, 11-13 (1980).

High-resolution polarization images of Crab Nebula with a charge-coupled device camera

Ian S. McLean*, C. Aspin† & H. Reitsema‡

* Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK

† Department of Astronomy, University of Edinburgh, Edinburgh EH9 3HJ, UK

‡ Department of Planetary and Space Sciences, University of Arizona, Tucson 85721, USA

Optical linear polarization of the Crab Nebula supernova remnant is presumed to be a good example of the astrophysical occurrence of synchrotron emission from relativistic electrons in a magnetic field¹⁻³. However, there have been few attempts to map the spatial distribution of the polarization at the highest possible resolution and with high signal-to-noise ratio, although such observations contain information about the geometry and the small-scale structure of the magnetic field. We describe here observations of the Crab Nebula using a new charge-coupled device (CCD) imaging/spectropolarimeter. Two large regions each 150 arc s across were observed with a seeing-limited spatial resolution of 1.5 arc s: the 'eastern bay' and the central wisps. The pulsar was included in both fields. In general, we find no evidence for substantial small-scale (1.5 arc s) structure in the projected magnetic field, but there are changes on scales of the order of 5 arc s which would be barely discernible by photographic methods. In addition, there is remarkable organization of the field over domains ranging in size from 10 to 30 arc s. The uncertainty about the nebular polarization component near the pulsar is resolved. The presence or absence of small-scale structure in the field is important in understanding the energy transport mechanism for the relativistic electrons, and little is known about the connection between small polarization features and the remarkable filamentary structure in the nebula. Photographic maps² had a resolution of only 5.25 arc s and rather limited polarimetric accuracy; more recent photo-electric measurements⁴ yielded 1×3 arc s resolution but only in a small patch. Instrumental technique is crucial.

Linear polarization maps of the Crab Nebula were obtained on 23 November 1981 using the Royal Observatory, Edinburgh

† Present address: Ball Aerospace Systems Division, Boulder, Colorado 80306, USA.

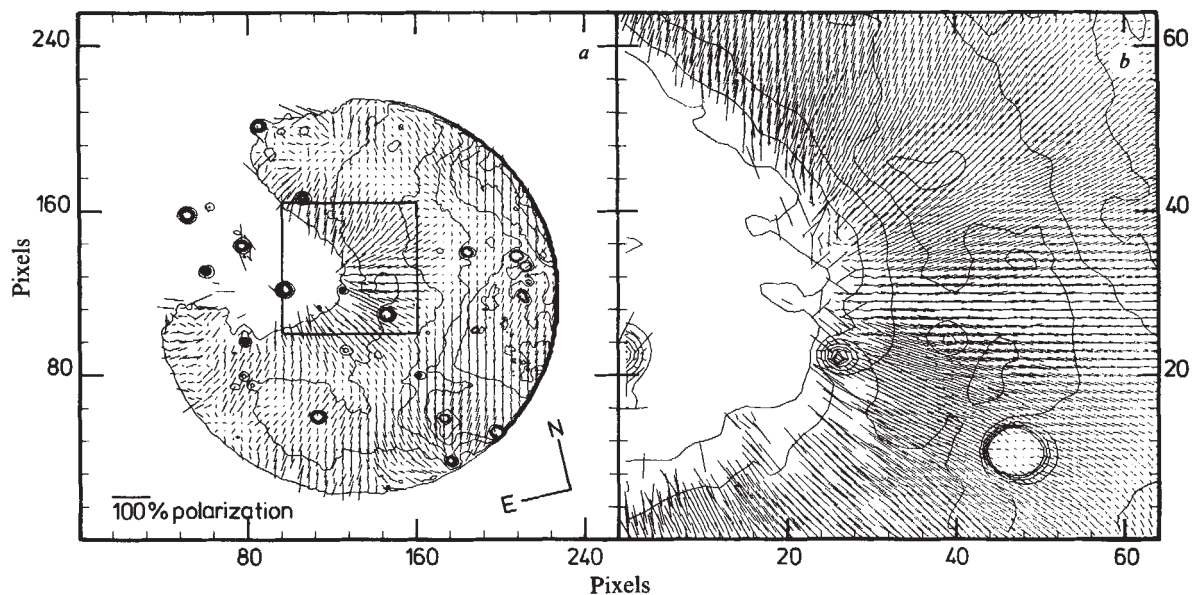


Fig. 1 *a*, Smoothed low-resolution (~ 4 arcs) CCD linear polarization-vector map, at a wavelength of 750 nm, overlaid on an intensity contour plot of the eastern bay of the Crab Nebula. The median polarizations in a 5×5 array on spacings of 5 pixels are displayed. The field of view is limited by a 150 arc s circular diaphragm. A polarization of 100% is represented by 12 pixels. *b*, Full resolution (0.75 arcs per pixel) polarization vector map of the indicated sub-section of *a*. The contour interval has been halved and a small component of interstellar polarization (that is, 2.0% at 160°) has been removed. Six pixels represent 100% p .

CCD imaging/spectropolarimeter (ISP) system mounted on the University of Arizona Observatories 1.54-m telescope on Mt Bigelow. The ISP and CCD camera are described in detail elsewhere⁵. Briefly, the polarimeter section comprises a rotatable half-wave plate located just above the focal plane entrance aperture followed by a Glan polarizing prism below the aperture. (A double beam prism using a calcite block is also available for smaller fields of view or when weather conditions are non-photometric.) Following the polarimeter section is a filter photometer containing a collimating lens, a filter wheel and a fast camera lens to re-image the focal plane onto the CCD. With the 1.54-m telescope the final scale on the CCD was 0.75 arcs per pixel. For these measurements, a bulk silicon RCA CCD having 30- μm pixels on 30- μm centres in a 512×320 array format was used. The CCD was cooled to -130°C . Because of vignetting restrictions imposed by the size of the polarizing optics, the field of view was limited to 150 arc s with a circular diaphragm. In this 'wide-field' (single-beam) mode, a sequence of four CCD frames corresponding to four half-wave plate positions differing in orientation by $22.5^\circ \pm 0.25^\circ$ are required to produce maps of the total intensity (I) and the Stokes parameters (Q , U) of the linear polarization; $Q = p \cos 2\theta$ and $U = p \sin 2\theta$ where p is the total fractional linear polarization and θ is the position angle of the direction of vibration of the electric vector measured eastwards from north. The Q frame is formed from the normalized difference of two halfwave plate positions at $a_0 + 0^\circ$ and $a_0 + 45^\circ$ (a_0 is arbitrary); the U frame is similar but the settings are $a_0 + 22.5^\circ$, and $a_0 + 67.5^\circ$. a_0 is determined from observations of standard stars.

Measurements were made in continuum light from the nebula using an interference filter with a central wavelength of 7,500 Å and a passband of 800 Å (FWHM). This choice exploits the excellent red sensitivity of the CCD and avoids contamination by strong optical emission lines. By comparing polarization in the red with that observed in the blue/green one can immediately test the prediction of the synchrotron model that p be independent of λ . Two overlapping areas of the Crab Nebula, each 150 arc s in diameter and containing $\sim 30,000$ resolution elements, were chosen. The first area was centred on the eastern bay region ~ 1 arc min east of the pulsar, while the second region was located about 20 arc s west of the pulsar to encompass the central nebula and wisps. Each exposure was 12 min;

the entire sequence of eight exposures (two sets of four), plus four flat fields taken immediately after on the inside of the dome at the same orientation of the telescope, was accomplished in < 1 h 45 min. For exposures of 12 min the photon noise on the signal (in one pixel) from the night sky was comparable with the readout noise (~ 150 e for this particular CCD; much lower noise devices have since been obtained). Our observations were made as the Crab Nebula passed overhead under photometric seeing conditions—stellar images measured 1.5 arc s FWHM on all frames. Subsequent data processing showed that, on average, registration of frames was good to better than 0.5 pixel (0.4 arc s) over the entire observation time. Using software developed for the Starlink computer node at the Royal Observatory, Edinburgh the frames were ultimately flat-fielded (that is corrected for pixel to pixel variations in sensitivity) and then registered to 0.04 pixels (0.03 arc s) by two-dimensional gaussian fits to stellar images. The 1σ errors in polarization computed both from photon-counting statistics and small box averages on the Stokes parameters themselves were typically in the range 0.5–2.0% over most of the nebula.

All data have been corrected for the background night sky and for the systematic increase in p at low signal-to-noise ratios, that is we form $p^2 = Q^2 + U^2 - \sigma^2$, where σ^2 is the quadrature sum of the errors in Q and U . A threshold level corresponding to about 2 magnitudes fainter than the night sky has been adopted as a cutoff point below which the signal from the nebula is deemed too faint to yield reliable polarizations. Instrumental polarization is negligible. Of course, the CCD does contain some bad individual pixels which can yield spurious polarizations. These are easily identifiable, especially in the regions of overlap. We have interpolated across only the most obvious defect, a blocked column, and have avoided an extensive treatment. The measurements are remarkably consistent throughout the region of overlap between the two fields. For example, the average polarization of 50 pixels 5 arc s north-west of the pulsar yielded $14.2 \pm 0.5\%$ at $148^\circ \pm 2^\circ$ in the eastern bay field, and $13.3 \pm 0.5\%$ at $147^\circ \pm 2^\circ$ in the central field.

Our new observations in the R band compare very well with previously published measurements at bluer wavelengths, such as the photographic maps by Woltjer² with 5.25 arc s diameter resolution, the single-channel, photoelectric data of Hiltner⁶

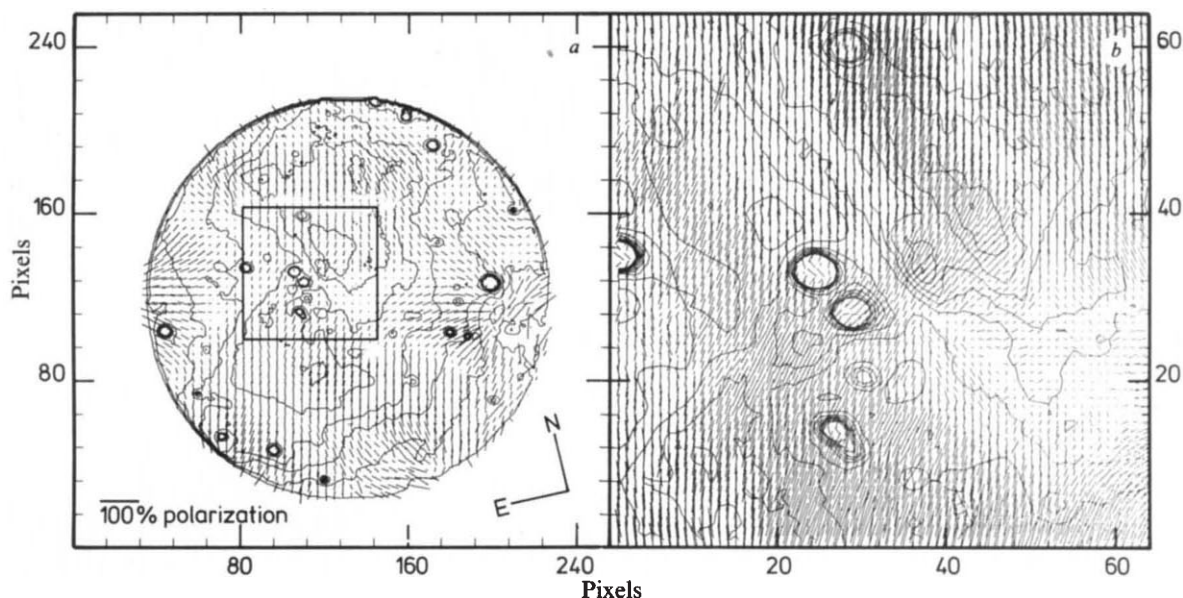


Fig. 2 *a, b*, As for Fig. 1 but for the central regions of the Crab Nebula containing the pulsar and wisps. The scale length for 100% p vectors is 12 pixels. The pulsar is almost central at pixel coordinates (28, 28); the 'anvil' is the region of uniform polarization extending south-east of the pulsar (20, 22), and the 'wisps' are the bright regions to the north-west around (38, 36). The 1950 coordinates of the pulsar are 5 h 31 min 31.5 s +21°58'55".

with a 5 arc s aperture and the Digicon strips by Schmidt *et al.*⁴ with 1×3 arc s or 3×3 arc s resolution in two small rectangular areas. For example, 30 arc s north-west of the pulsar we find $24 \pm 1\%$ at $171^\circ \pm 1.5^\circ$ in a 5 arc s patch while Schmidt *et al.* report 23% at 168° and Woltjer's map indicates 30% at 171° . Schmidt has noted that the photographic data tend to overestimate p . Of course, there are very few photoelectric measurements at red wavelengths. Gehrels' reports 22% at 160° for a region 30 arc s in diameter lying 20 arc s south-west of the pulsar ($\sim 53, 13$ in Fig. 2*b*). Our map shows a huge range in p and θ values in this region but the mean values agree well (21.5% at 159.8°). Our observations are also in accord with photoelectric measurements of the pulsar itself. From 4 pixels (two per region) covering the pulsar we obtain an observed time-averaged polarization of $5.5 \pm 0.8\%$ at $103^\circ \pm 4^\circ$ which agrees with older published values^{4,8,9} of $\sim 6.5\%$ and angles between 98° and 107° , and with the high resolution pulse profiles by Jones *et al.*¹⁰.

Figures 1*a* and 2*a* show intensity contour plots in the R filter overlaid by a smoothed, low spatial resolution polarization-vector map for each region of the nebula. Note that only 1/25 of the available CCD data are displayed in these smoothed maps but even so the resolution exceeds that of the old photographic maps. Figures 1*b* and 2*b* show two smaller regions of the nebula at maximum spatial resolution and include the region of the pulsar, wisps and anvil, and the well-known eastern bay region. The polarization shown in Figs 1*a* and 2*a* is the observed polarization. Since the nebula lies at a relatively low galactic latitude the observations suffer the superposition of an interstellar component. Following previous authors¹¹ we adopt $p = 2.0\%$, $\theta = 160.0^\circ$ as the best estimate of the interstellar polarization. This small component has been removed for the higher resolution plots of Figs 1*b* and 2*b* which therefore display the intrinsic polarization.

Of course, the maps show only the projection onto the plane of the sky of the polarization integrated along the line of sight from the back to the front of the nebula. In parts of the nebula, such as the eastern bay, this integration appears to be minimal and isolated magnetic field loops are apparently observed. Consequently the polarization can become large and approach the values expected from synchrotron theory for isotropic pitch angle distribution which is $p = (\gamma + 1)/(\gamma + 7/3)$ where the spec-

tral index¹² γ is of the order of 2.0 (that is $p_{\max} \sim 69\%$). An important result from our study is that large polarizations (50% or more) occur only around the periphery of the nebula (compare with the eastern bay) but not in the main body of the nebula even at high spatial resolution.

We have searched for isolated polarization knots and high polarizations associated with intensity features seen in images in both continuum light and in emission lines¹³. In particular, we sought evidence of field lines curving around or being disrupted by filaments. It is not obvious how or whether the observed polarization patterns are connected with such features. There is, however, a strong suggestion that the troughs of low polarization which traverse the nebula may correlate with emission filaments having negative radial velocities. These regions may represent tangled fields or they may be boundaries (towards the front of the nebula) between large well-ordered magnetic loops.

Down to the seeing limit of ~ 1.5 arc s (which corresponds roughly to 15 light days (0.01 pc) in the nebula, the extent of which is about 10 light yr (3 pc) across the major axis) there is little or no detailed structure in the continuum polarization. This lack of structure is a significant result for diffusion models of energy transport which might be expected to show structure on a scale of 3–5 arc s. On the other hand, it is clear from the images in Figs 1 and 2 that the polarization pattern often maintains a remarkably high degree of uniformity over 'domains' of typically 20 arc s. Each domain is separated from an adjacent one by a transition region in which the polarization changes smoothly on a scale of ~ 5 arc s. Within any given domain the polarization remains remarkably stable, typically with a value in the range of 25–40%. Sometimes the difference between domains is mainly that the net orientation of the E -vectors is different. Close examination of the earlier published observations^{2,4} reveals that the same effect is present, but it is clearly much more apparent at higher spatial resolution.

One explanation of the regular domains is that the magnetic field in the nebula is relatively simple and uniform but is distorted by different amounts in different directions—presumably by interaction with the interstellar medium. Along any given line of sight one sees polarized radiation from several discrete loops¹⁴ or bubbles located one behind the other. Integration of the polarization vectors along such a line of sight

Table 1 Average polarization in 0.75 arcs wide box annuli centred on the Crab pulsar

Annulus no.	No. of pixels	Mean intensity	Mean $P_{\text{obs}}(\%)$	Mean θ_{obs}	θ_{NEB}
Pulsar	2	4,390	5.5 ± 0.8	$103 \pm 4^\circ$	
1(0.75)	10	$2,550 \pm 170$	11.8 ± 1.1	132 ± 3	138
2(1.50)	18	$1,560 \pm 45$	15.1 ± 0.4	138 ± 2	140
3(2.25)	26	$1,380 \pm 26$	14.0 ± 0.3	144 ± 1	145
4(3.0)	31	$1,317 \pm 26$	12.8 ± 0.4	147 ± 1	147
5(3.75)	37	$1,299 \pm 23$	12.6 ± 0.4	150 ± 1	150
6(4.25)	43	$1,287 \pm 21$	13.1 ± 0.4	150 ± 1	150

The number in parenthesis is the inner radius of annulus in arcs. θ_{NEB} is position angle in nebula corrected for pulsar contamination. The nebular background corresponds to $\sim 2,311$ units per arc s² or ~ 19.75 mag in R.

will tend to reduce the individual highly polarized contributions. Suppose there are n such loops, and that the volume emissivity of synchrotron radiation is uniform throughout the nebula. The latter condition is supported by the smoothness of the polarization maps. Then in the ideal (random walk) case one might expect a depolarization factor of the order of $1/\sqrt{n}$. Assuming $p_{\text{max}} \sim 69\%$ then an observed polarization of $p \sim 35\%$ corresponds to only four loops, 23% corresponds to nine loops—a remarkably simple geometry.

Studies of the size of the emission region at radio, optical and X-ray wavelengths have also been used to examine energy transport models^{15,16}. Diffusion models are generally found to be less satisfactory, at least for the radio and optical emitting electrons, than bulk motion models such as that by Rees and Gunn³. The new high-resolution polarization maps are consistent with the idea that power from the pulsar emerges partly as a 30-Hz wave and partly as a relativistic wind containing a toroidal field. Away from filamentary structure the magnetic field could be expected to be fairly well ordered—just as we observe. If the wisps represent the shock front encountered by the 30-Hz wave then one might expect the magnetic field to be aligned with (the electric vector perpendicular to) the wisps. This was found to be the case by Schmidt *et al.* when allowance is made for the foreground nebular contribution to the polarization and our data confirm their conclusion.

It was generally believed that observations of the pulsar during the interpulse period would yield a reasonable value for the nebular polarization component. Kristian *et al.*⁸ reported a position angle of $\sim 159^\circ$ in disagreement with Cocke *et al.*¹⁷ who obtained only 138° . Kristian *et al.* claimed that the polarization was the same in apertures of 5, 12 and 19 arc s, whereas Cocke *et al.* used an aperture of 3 arc s. Our high-resolution data reveal that unless small apertures are used the results can be confusing. Close to the pulsar the nebular polarization is quite uniform (at ~ 12 – 13%) but the position angle changes steadily with radial distance. Table 1 shows the mean position angles in box-like annuli each 0.75 arc s wide surrounding the pulsar; 2–3 arc s from the pulsar the mean value is around 141° but beyond 5 arc s the position angle exceeds 155° and is very direction-dependent. Disney¹⁸ has shown that a value of 140° is consistent with the rotating perpendicular vector model of the pulsar.

Jones *et al.*⁹ have recently demonstrated that the interpulse polarization of the Crab pulsar is much higher than previously thought, indicating an origin close to the speed-of-light cylinder. Their technique relied on adopting the background polarization as that within an annulus only 1.3 arc s wide with a mean radius of 2.0 arc s centred on the pulsar. Figure 2b indicates that the nebular background is sufficiently smooth to justify their assumption. Finally, these authors report a large disagreement in p , θ (seen in their small aperture) with a region 50 arc s north-west of the pulsar quoted by Schmidt *et al.* which they interpret as small scale structure. In fact, the Schmidt paper contains a misprint (G. D. Schmidt, personal communication) and the intended region is only 20 arc s from the pulsar. Our

results agree with the findings from both regions and dispel the suggestion of small scale structure.

A further analysis of these new polarization maps is in progress. Many other previously non-viable observations are now possible with this instrument and its CCD detector.

We thank the University of Arizona Observatories for allocation of telescope time and our many colleagues at the Royal Observatory, Edinburgh for their advice and support. We also thank Martin Rees, F. Graham Smith and Mike Disney for comments. This work was supported by SERC.

Received 22 March; accepted 23 May 1983.

- Shklovskij, I. S. *Cosmic Radio Waves* (transl. Rodman, R. B. & Varsavsky, D. M.) (Harvard University Press, 1960).
- Woltjer, L. *Bull. astr. Inst., Netherlands* **13**, 301–311 (1957).
- Rees, M. J. & Gunn, J. E. *Mon. Not. R. astr. Soc.* **167**, 1–12 (1974).
- Schmidt, G. D., Angel, J. R. P. & Beaver, E. A. *Astrophys. J.* **227**, 106–113 (1979).
- McLean, I. S., Cormack, W. A., Herd, J. T. & Aspin, C. *Soc. Phot. Inst. Engng.* **290**, 155–158 (1980).
- Hiltner, W. A. *Astrophys. J.* **125**, 300 (1957).
- Gehrels, T. *Planets, Stars and Nebulae Studied with Photopolarimetry*, 1028 (ed. Gehrels, T.) (University of Arizona Press, Tucson, 1974).
- Kristian, J., Visvanathan, N., Westphal, J. A. & Snellen, G. H. *Astrophys. J.* **162**, 475–483 (1970).
- Wampler, E. J., Scargle, J. D. & Miller, J. S. *Astrophys. J. Lett* **157**, L1–L10 (1969).
- Jones, D. H. P., Smith, F. G. & Wallace, P. T. *Mon. Not. R. astr. Soc.* **196**, 943–953 (1981).
- Martin, P. G. & Angel, J. R. P. *Astrophys. J.* **193**, 343–351 (1974).
- Baldwin, J. R. *The Crab Nebula*, IAU Symp. No. 46, 22–31 (1971).
- Scargle, J. D. *Astrophys. J.* **156**, 401–426 (1969).
- Felton, J. E. in *Planets, Stars and Nebulae Studied with Photopolarimetry* (ed. Gehrels, T.) 1014–1028 (University of Arizona Press, 1974).
- Wolff, R. S. & Novick, R. *Astrophys. Lett.* **18**, 29–33 (1976).
- Adams, D. J., Eaton, N. C. & Lawson, P. A. *Observatory* **103**, 20–24 (1983).
- Cocke, W. J., Disney, M. J., Muncaster, G. W. & Gehrels, T. *Nature* **227**, 1327–1329 (1969).
- Disney, M. J. *Astrophys. Lett* **9**, 9–12 (1971).

A waveguide for low-frequency electromagnetic waves in the upper ionosphere

A. B. Shvartsburg

IZMIRAN Troitsk, p/o 142092, Moscow Region, USSR

New possibilities for extremely low frequency (ELF) radio communication are shown here to arise from the peculiar waveguide properties of the anisotropic multicomponent magnetoplasma in the upper ionosphere. Unlike the well known high-frequency (HF) communication channels formed by the electron density distribution in the lower ionosphere¹, the variable ionic composition and the geomagnetic field are essential for this spectral range below the ionic gyrofrequency. This global waveguide cavity, effective for Alfvén wave propagation, surrounds the Earth near the ionospheric F₂-layer maximum. Unlike the LF waveguide cavity, located between the Earth's surface and the lower ionosphere², the aforesaid cavity is not connected with the boundaries of any media: the waveguide-like distribution of refractive index for Alfvén waves is determined by smooth altitudinal profiles of electron and ionic densities. The dispersive properties of these waveguides are very sensitive to the proportions of heavy and light ions; therefore, they may be termed 'chemical' waveguides.

Let us consider the propagation of ELF extraordinary Alfvén waves. The refractive index η of such waves is isotropic and may be written in the collisionless approximation as³

$$\eta^2 = \varepsilon_{\perp} \quad (1)$$

$$\varepsilon_{\perp} = \frac{\omega_p^2}{\omega_H^2} \sum_{\alpha} \frac{\gamma_{\alpha}}{\beta_{\alpha}} \quad \gamma_{\alpha} = \frac{N_{\alpha}}{N_e}; \quad \beta_{\alpha} = \frac{m}{M_{\alpha}} \quad (2)$$

Here ω_p and ω_H are, respectively, the values of plasma frequency and gyrofrequency of electrons; N_e and N_{α} are the densities of electrons and ions of species α ; m and M_{α} are

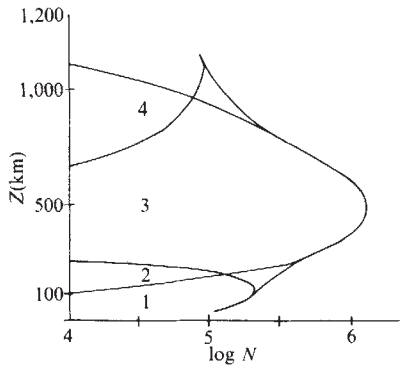


Fig. 1 Altitudinal profiles of electron density⁴, N_e , and ionic composition over the magnetic equator; curves 1, 2, 3 and 4 refer to the electron density and to the densities of molecular ions, H^+ and O^+ , respectively.

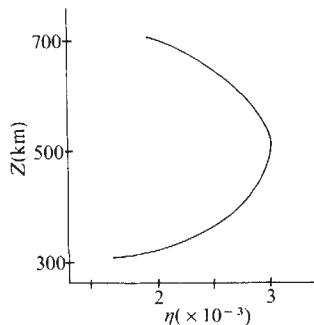


Fig. 2 Altitudinal distribution of the refractive index η of extraordinary Alfvén waves, responsible for the chemical waveguide for the propagation of these waves.

electron and ion masses; the summation in equation (2) is carried out for all types of ions. We consider the frequencies ω , restricted by the condition

$$\omega^2 \ll \omega_H^2 \beta_\alpha^2 \quad (3)$$

in a relatively dense plasma, typical for the upper ionosphere, for which

$$\omega_p^2 \gg \omega_H^2 \quad (4)$$

The space-time variations of electron density $N_e(z)$ and the ionic composition $N_\alpha(z)$ given in equation (2) may change the location and properties of these LF waveguide channels. The main species of ions in these regions are O^+ , H^+ and NO^+ . The values of $N_\alpha(z)$ over the magnetic equator in quiet solar conditions are plotted in Fig. 1.

The altitudinal variation of refractive index η , given by equation (1), using the electron density distribution $N_e(z)$, obtained by means of the incoherent scatter radar at Jicamarca⁴, is illustrated in Fig. 2; the frequency shown satisfies the inequality (4). This graph indicates the existence of ELF waveguide in the upper ionosphere, above the F₂-layer maximum, $h_m = 450$ km at heights between 500 and 600 km. The wavelengths for these frequencies ($\omega \leq 60$ rad s⁻¹) are restricted by the extremely large values of refractive index, $\eta \approx (2+3) \times 10^3$, to $\lambda \geq 12-15$ km. Near its maximum, $Z = Z_0$, the refraction index variation, shown in Fig. 2, may be approximated by

$$\eta_0^2 \left[1 - \frac{(Z - Z_0)^2}{2L^2} \right]; \quad (5)$$

Here $\eta_0 = \eta(Z_0)$ is the maximum value of η at height Z_0 . This distribution, modified gradually in longitudinal and altitudinal directions, forms the waveguide cavity, limited in the z -direction, for extraordinary Alfvén waves. The spectrum of

waveguide modes, propagated both in latitudinal and longitudinal directions, may be analysed by means of a dimensionless scalar wave equation for the function h_θ , which governs the longitudinal component of the variable magnetic field of Alfvén's wave h_θ :

$$\frac{\partial^2 h_\theta}{\partial X^2} + (K_\perp^2 V^2 - X^2) h_\theta = 0;$$

$$X = V(Z - Z_0); \quad V^2 = \frac{cL}{\sqrt{2}\omega\eta_0} \quad (6)$$

here $K_\perp$ is the vertical component of the wave vector.

Equation (6) is well known in the theory of the quantum oscillator⁵. The wavefield is localized in the z -direction near $Z = Z_0$ if the values of the component $K_\perp$ are given by

$$K_{\perp m}^2 = \frac{\sqrt{2}\omega\eta_0}{cL} (1 + 2m); \quad m = 0, 1, 2, \dots \quad (7)$$

Thus, the spectrum of $K_\perp$ is quantized. We may evaluate the thickness of the waveguide layer ρ_c , associated with the fundamental mode of propagation, by means of the formula⁵:

$$\rho_c = \sqrt{\frac{2L^2 c^2}{\omega^2 \eta_0^2}} \quad (8)$$

The substitution of the data, illustrated in Fig. 2 ($\omega \leq 60$ rad s⁻¹), into this formula leads to the value $\rho_c \approx 20$ km. The waveguide regime of propagation inside this cavity is characterized by the restricted number of Eigen modes from equation (7): $m \leq 10$.

The attenuation of such waves may be determined by means of the set of equations, which governs the motion of ions, electrons and neutral species³. The standard procedure of calculating the complex refractive index ($\eta - i\mathcal{K}$) permits us to obtain the real part η in a form of equation (1)–equation (2), and the attenuation coefficient $K_\mathcal{K}$ depending on the frequencies of electron collisions with heavy particles ν_e , the ion–neutral species collisions $(\nu_i)_\alpha$ and the ratio $N_e N_m^{-1}$, N_m being the density of neutral species at the height $Z = Z_0$. The ratio N_e/N_m is small here⁶ $N_e/N_m \approx 3-5 \times 10^{-2}$ and the frequencies ν_e and $(\nu_i)_\alpha$ satisfy the condition $\beta_2 \nu_e \gg (\nu_i)_\alpha$ for each species of ion α . The attenuation coefficient $K_\mathcal{K}$ in such conditions may be written as

$$K_\mathcal{K} = \frac{\omega}{c} \mathcal{K} = \frac{\eta_0}{2c} \sum_\alpha \frac{\gamma_\alpha}{\beta_\alpha} (\nu_i)_\alpha \quad (9)$$

The greatest contributions to equations (2) and (9) calculated for the 'chemical' waveguide regions, are from O^+ and H^+ . The dissipation length $L_\mathcal{K}$

$$L_\mathcal{K} = K_\mathcal{K}^{-1} \quad (10)$$

is large ($4-7 \times 10^3$ km). The collisional dissipation of these waves in the lower ionosphere would be considerable; thus, the excitation of such waveguides by means of satellite antennae seems to be more expedient.

Such a global waveguide cavity may be effective for Alfvén waves ($\omega \leq 60$ rad s⁻¹), propagated in the arbitrary directions, including meridional and latitudinal ones. These waves, transmitted from a satellite, may be used for radio communication, navigation systems or for remote ionospheric sounding.

Received 1 November 1982; accepted 18 May 1983.

1. Pitteway, M. L. & Cohen, R. *J. geophys. Res.* **66**, 3141 (1961).
2. Schumann, W. *Z. Angew. Phys.* **9**, 373 (1957).
3. Ginzburg, V. L. & Rukhadze, A. A. *Waves in Magnetoplasma* (Nauka, Moscow, 1975).
4. Clark, W. L., McClure, J. P. & Van Zandt, T. E. *Rep. VAG-53* (World Data Center A, Boulder, 1976).
5. Landau, L. D. & Lifshitz, E. M. *Quantum Mechanics* (Nauka, Moscow, 1963).
6. Jaccia, L. G. *Thermospheric Temperature, Density and Composition: New Models* (SAO Spec. Rep. no. 375, 1977).

Laser radar observations in mid-Wales of aerosols from the El Chichón eruption

L. Thomas, D. B. Jenkins, D. P. Wareing & M. Farrington

Department of Physics, University College of Wales, Aberystwyth, Dyfed SY23 3BZ, UK

Observations using a laser radar system at Aberystwyth ($52^{\circ}25' \text{N}$, $4^{\circ}4' \text{W}$) have shown enhanced concentrations of aerosols at heights up to $\sim 30 \text{ km}$. It is believed that these enhancements have been caused by eruptions of El Chichón ($17^{\circ}33' \text{N}$, $93^{\circ}20' \text{W}$) in southern Mexico during the period 28 March to 4 April 1982. Data collected during 45 nights in the period November 1982 to February 1983 have shown the presence of such enhancements consistently since observations started on 6 November and represent the most northern excursion of the stratospheric cloud by that date^{1,2}. The results indicate a very pronounced variability in the cloud, both on a night-to-night basis and within given nights, the peak value for the ratio of backscatter from aerosols to that from molecules varying between 2 and 12. In these changes the aerosols can serve as a tracer of transport processes at stratospheric heights³.

The radar system incorporates a Nd/YAG laser operating at the second harmonic wavelength of 530 nm with 15-ns pulses at a repetition frequency of 10 Hz. The receiver includes a

1.0-m newtonian telescope and is operated in a photon counting mode, the backscattered signals being recorded in channels of widths selected according to the desired height resolution. The incorporation of a suitable interference filter and neutral density filters, together with a synchronized motor-driven aperture shutter, ensures background light rejection and minimizes the errors associated with pulse overlap and photomultiplier overloading by returns from low altitudes.

Results obtained with a 300-m height resolution over the 4-month period of observation are illustrated in Fig. 1. This shows the variation with height of the range-corrected backscattered signal for about 40 min of observations on each of 12 selected nights. In each case the height variation expected for Rayleigh scattering, as computed for molecular densities from an atmospheric model for 45°N latitude in midwinter⁴, is shown by a broken curve; no data on molecular densities are available for individual nights over the entire height range of the present measurements. The observations on each night showed an approach to the Rayleigh-scatter profile at heights above 30 km; this profile was therefore fitted to the observations at these heights and extrapolated down to $\sim 15 \text{ km}$. The standard errors of the measurements are shown at intervals of 3 km at the greater heights, these errors becoming negligible with decreasing height. The abscissa scale allows estimates to be made of the ratio of aerosol-to-molecular backscatter at the greater heights. The results show a marked enhancement of stratospheric aerosols when compared with previous measurements in the United Kingdom⁵.

The early observations in Fig. 1 show an increase in the aerosol-to-molecular backscatter ratio from about 2 on the first night of observation, 6 November 1982, to a peak value of about 10 near 22 km on 19 November 1982, with a partial

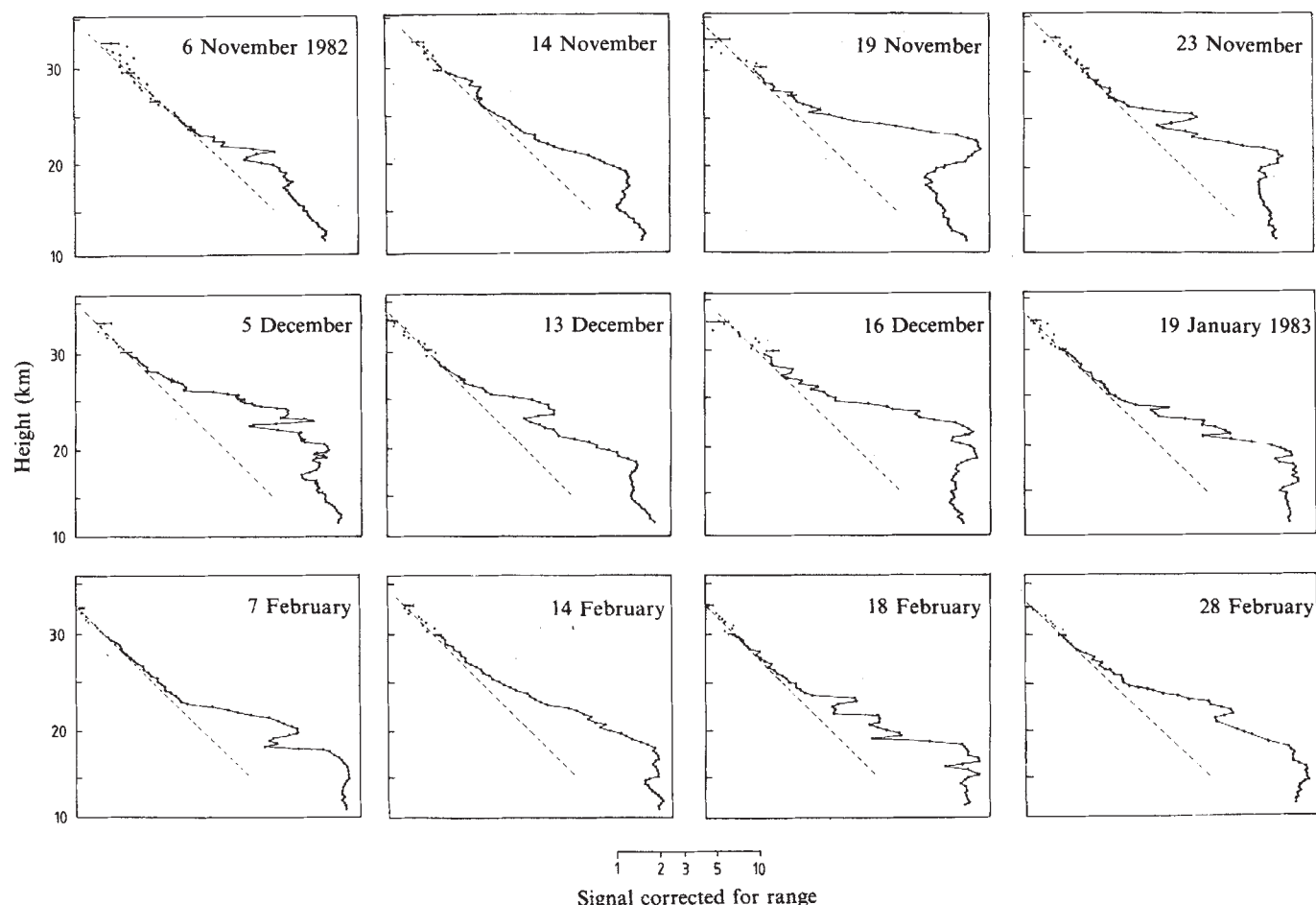


Fig. 1 Variations of the range-corrected signals from soundings at vertical incidence on selected evenings during November 1982–February 1983. The broken curves represent the corresponding variation expected for Rayleigh scattering as computed for molecular densities derived from an appropriate atmospheric model⁴.

recovery on 23 November but with a second layer being shown near 25 km. The results for 5 December 1982 show a more striking vertical structure in the aerosol layer and a change in form occurred between 13 and 16 December. The observations on the four nights in February 1983 further demonstrate the marked variability in the layer.

The signals shown at the lowest heights in Fig. 1 will have been influenced by scattering from aerosols and, perhaps, ice crystals at these heights and also attenuation at greater heights due to Rayleigh scattering and scattering⁶/absorption by aerosols. The magnitude of these latter effects due to aerosols are difficult to assess in the absence of information on the size distribution and optical properties of the aerosols but would be relevant to the observations of increased temperatures in the equatorial lower stratosphere which have been attributed, at least in part, to aerosols from the eruption^{2,7}.

In the absence of data for individual nights, the molecular densities from various models⁴ have been used to examine the range of variability that might be expected in the Rayleigh scatter signals. It is considered that the use of a common Rayleigh scatter profile, and also the attenuation at greater

heights referred to above, introduce no significant errors in the estimates of the maximum aerosol-to-molecular scatter ratio and contribute little to the night-to-night variability indicated in Fig. 1.

This variability was often shown in the results for successive nights, and even within shorter periods. This is illustrated in the change between the profiles representing 25 min of observation centred at 0646 UT and 1742 UT on 27 November 1982 in Fig. 2; this change corresponds to an increase in the peak value of the aerosol to molecular backscatter ratio from about 4 to 12. The second pair of profiles included in Fig. 2 refer to similar times on the previous day, 26 November 1982; although a change in form is observed, no pronounced change in the backscatter ratio is shown.

While the changes observed in Figs 1 and 2 probably result from horizontal movements of the stratospheric aerosols, predominantly in the zonal direction, certain features of the short-term changes indicate apparent vertical motions. An example is shown in the sequence of profiles each observed over periods of about 25 min at approximately 2-h intervals on the evening of 24 November and the early morning of 25 November

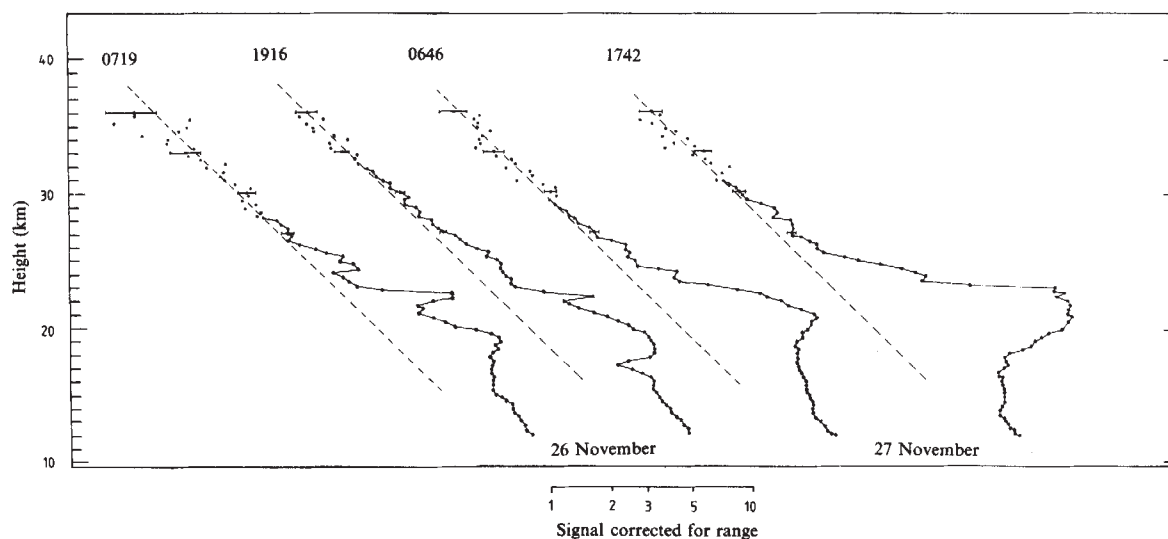


Fig. 2 Variations of the range-corrected signals from soundings at vertical incidence before sunrise and following sunset on 26 and 27 November 1982; the times shown are UT. The broken curves represent the corresponding variation expected for Rayleigh scattering as computed for molecular densities derived from an appropriate atmospheric model⁴. Attention is drawn to the pronounced change between the two profiles on 27 November, this change not being shown by the corresponding data for 26 November.

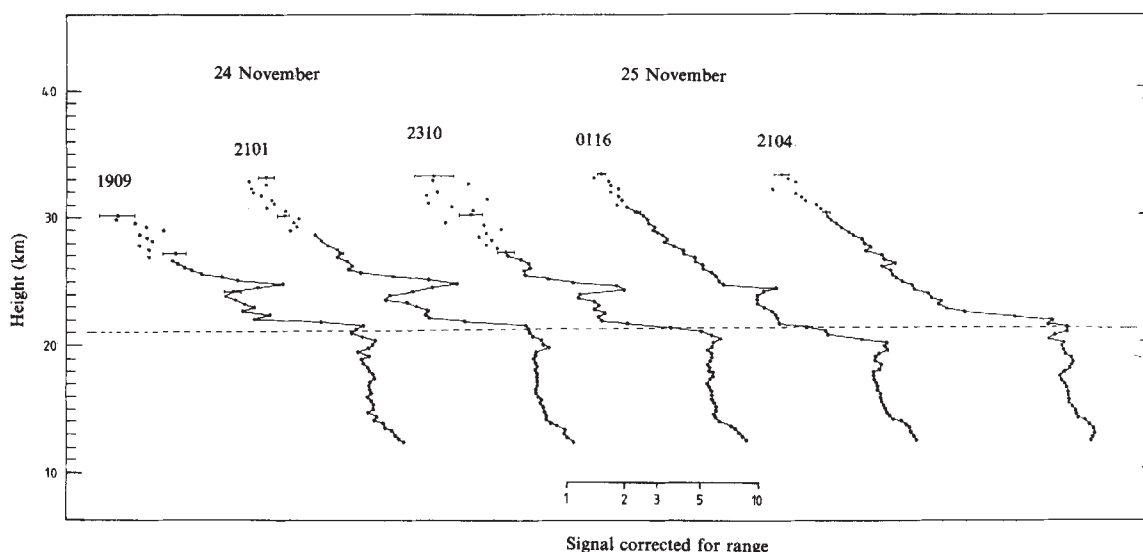


Fig. 3 Variations of the range-corrected signals from soundings at vertical incidence during the evening of 24 November, and the early morning and evening of 25 November 1982. Attention is drawn to the successive lowering of the ledge near 21 km in the backscatter profile and in the subsidiary maximum near 24 km between 1909 UT on 24 November and 0116 UT on 25 November. The ledge had returned to its initial height by 2104 UT on 25 November.

(Fig. 3). This sequence shows a lowering of the main ledge in the profile, near 21 km, and of the peak of the subsidiary maximum near 24 km, over a height range of about 1.5 km in 5 h, corresponding to an apparent vertical velocity of $\sim 0.08 \text{ m s}^{-1}$. The subsequent changes could not be followed because of obscuration by low-altitude cloud but the main ledge had returned to its initial height by 2104 UT on 25 November. The velocity of 0.08 m s^{-1} is too large to be reconciled with either vertical winds⁸ or gravitational settling⁹ and the sequence shown in Fig. 3 probably represents the horizontal motion of a tilted structure. With zonal mean winds of $\sim 10 \text{ m s}^{-1}$ expected near 20 km altitude at these latitudes in November¹⁰, the changes observed on 27 November (Fig. 2) and 24–25 November (Fig. 3) would correspond to horizontal dimensions of $\sim 400 \text{ km}$ and 200 km , respectively.

The general variability in the profiles of backscatter ratio serves to obscure any long-term change after mid-November. However, the observations over the 4-month period show little evidence of a decay in the aerosol concentration; on the basis of previous studies¹¹, a decrease by a factor of about 0.6 might have been expected in the absence of a continuing arrival of new material. Such a decay will only become obvious when the major redistribution of the stratospheric aerosols has been completed.

Received 25 April; accepted 25 May 1983.

1. EOS 64, 1 (1983).
2. Labitzke, K., Narjokat, B. & McCormick, M. P. *Geophys. Res. Lett.* **10**, 24–26 (1983).
3. McCormick, M. P. in *Atmospheric Effects and Potential Climatic Impact of the 1980 Eruption of Mount St Helens* (ed. Deepak, A.) (125–130 NASA Conf. Publ. 1982).
4. U.S. Standard Atmosphere Supplements (U.S. Government Printing Office, Washington DC, 1966).
5. Thomas, L., Chaloner, C. P. & Bhattacharyya, S. K. *Nature* **289**, 473 (1982).
6. Hirono, M. & Shibata, T. *Geophys. Res. Lett.* **10**, 152–154 (1983).
7. Parker, D. E. & Brownson, J. L. *Nature* **301**, 406–408 (1983).
8. Dunkerton, T. J. *atmos. Sci.* **35**, 2325–2333 (1978).
9. Huntten, D. M. *Geophys. Res. Lett.* **2**, 26–28 (1975).
10. Newell, R. E., Kidson, J. W., Vincent, D. G. & Baer, G. J. *The General Circulation of the Tropical Atmosphere* Vol. 1 (MIT Press, Boston, 1974).
11. Hofmann, D. J. & Rosen, J. M. *J. atmos. Sci.* **38**, 168–181 (1981).

Long-term thermo-mechanical properties of the continental lithosphere

Garry D. Karner*†‡, Michael S. Steckler* & Julian A. Thorne*†

*Lamont-Doherty Geological Observatory, and †Department of Geological Sciences, Columbia University, Palisades, New York 10964, USA

Simple thermo-elastic models successfully explain flexure observations from oceanic lithosphere and are consistent with predictions based on empirical yield stress envelopes for the deformation of olivine. The response of the continental lithosphere, in contrast, has frequently been modelled as a Maxwell viscoelastic solid. By plotting estimates of flexural rigidity against the age of the continental lithosphere at the time of loading, an estimate of which is defined by the radiometric age of the underlying basement, we report here that flexural rigidities for the continental lithosphere are compatible with oceanic flexure results. Analogous to the increase of rigidity with age away from a mid-oceanic ridge, continental rigidities appear to increase with age following a major thermal event. The high rigidities ($\sim 10^{32} \text{ dyn cm}$) associated with Archaean/Proterozoic terrains and modelling of plate deformation with yield-stress envelopes appropriate for the continents suggest that the long-term thermal behaviour of continental lithosphere is governed by a cooling plate model with a 200–250-km plate thickness.

Attempts have been made to predict the stratigraphy of intracratonic basins¹, foreland basins², and passive continental margins^{2–5} based on simple models of the thermo-mechanical properties of both oceanic and continental lithospheres. Recent studies of the deformation of the lithosphere by applied loads, such as seamounts^{6–8} and trenches⁹, suggest that the oceanic lithosphere can be represented as an elastic plate overlying a weak fluid. The plate thickness, T_e , or equivalently, the flexural rigidity, D , determines the shape and width of the deflection. The flexural rigidity is related to T_e by $D = ET_e^3/12(1 - \nu^2)$ where E is Young's modulus and ν is Poisson's ratio.

Watts⁶ showed that the flexural rigidity increases with the age of the oceanic lithosphere. Furthermore, the depth to the base of the elastic plate appears to follow the 450 °C isotherm obtained from the cooling plate model of Parsons and Sclater¹⁰. The deflection associated with a loading event is determined only by the age of the plate at the time of loading and is maintained throughout geological time. Walcott¹¹, however, after studying the response of the continental lithosphere to geological loads, such as ice sheets, salt lakes, and individual mountains, concluded that the lithosphere was a visco-elastic Maxwell solid on time scales of 1–1,000 Myr. A Maxwell solid is characterized by an initial, high flexural rigidity which decreases with time following loading. Deflections and stresses created in a Maxwell substance, therefore, are not maintained through geological time. It is important to reconcile the apparent contradiction between the rheologies proposed for the oceanic and continental lithospheres.

In Fig. 1, the combined data of 20 oceanic and 14 continental estimates of flexural rigidity have been plotted as a function of the thermal age of the lithosphere at the time of loading. These data are similar to those used by Watts *et al.*¹, but include two new continental rigidities determined by Karner and Watts¹². For oceanic lithosphere, the thermal age directly relates to the plate age¹⁰, which is obtained from correlatable sea floor magnetic anomalies. For the continental lithosphere, an estimate of the thermal age is obtained from the radiometric age of the continental basement being flexed. By subtracting the age of the load from this plate or radiometric age, we obtain an estimate for the thermal age of the lithosphere at the time of loading.

The method used to obtain continental basement thermal ages may, however, introduce some error. For example, the outcropping basement rocks near the Michigan Basin are of Grenvillian age (K/Ar date $950 \pm 150 \text{ Myr}$; Rb/Sr date $1,150 \pm 100 \text{ Myr}$) and have not been reset by the presumed thermal event that initiated Michigan Basin subsidence. In contrast, basement rocks recovered from the deep borehole in the Michigan Basin¹³ have been metamorphosed to greenschist facies¹⁴ (implying temperatures of 400–450 °C), the timing of which may be related to the resetting of rock magnetism (and hence palaeomagnetic poles)¹⁵ during the latest Precambrian to Cambrian (500–800 Myr)¹³. Note that if this metamorphic age is assumed to be the thermally-reset age of the basement, the Michigan Basin result (Fig. 1) would move parallel to the x axis onto the observed trend.

Uncertainties in the flexural rigidity may also cause errors in the y -axis position of the data point. In particular, Walcott¹¹ only determined minimum rigidity estimates for the Caribou mountains and Interior plains (open circles, Fig. 1) and the actual rigidities may be significantly greater. Rigidities obtained from continent-wide studies using the statistical relationship between topography and its associated gravity anomaly^{16–19} have not been included in Fig. 1. As demonstrated by Forsyth^{20,21}, if several geological provinces with varying flexural rigidities and/or loads with little topographic expression are analysed using this technique, then the calculated admittance may mimic that of a low-rigidity plate²¹.

The flexure data included in Fig. 1 have been obtained from a wide variety of loads (such as seamounts, trenches, aseismic ridges, and river deltas for the oceanic lithosphere, and foreland basins, ice-sheets, salt-lakes, and intracratonic sedimentary

‡ Present address: Bureau of Mineral Resources, PO Box 378, Canberra ACT 2601, Australia.

basins for the continental lithosphere), environments, and different basement dating techniques. The most significant result is that the oceanic and continental data sets become consistent when continental rigidities are plotted in this way (that is, rigidity versus lithospheric age at the time of loading). The resultant simple trend therefore, is rather remarkable and suggests that errors introduced in estimating flexural rigidity and the basement age at the time of loading may be minor. There is an increase in rigidity with plate age from the relatively small values associated with mid-ocean ridges, to higher values associated with ice and sediment loads on older lithosphere. Haxby *et al.*²² and Molnar and Tapponnier²³ also suggested that the flexural rigidity of the continental lithosphere may increase with lithospheric age. Haxby *et al.*²² based their result on detailed stratigraphical isopach maps in the Michigan Basin while Molnar and Tapponnier²³ attributed the variation of deformation style in Asia to the collision between a cold, rigid, Indian continental block and a younger, less rigid, Eurasian block.

Also included in Fig. 1 are continental rigidity estimates from ice loading studies. Since the rigidity values from these studies are close to estimates from long-term foreland basin flexure, we conclude that the rigidities obtained from glacial lakes and associated stranded beach dune systems appear to be sensitive only to the elastic component of the continental lithosphere. Furthermore, loads emplaced on similarly aged lithosphere for times ranging from 10^4 yr (ice loading) to $> 10^8$ yr (Appalachian foreland) have the same effective rigidity (Fig. 1) suggesting that the rigidity does not depend on the time since loading.

The flexural rigidity estimates from the oceans are consistent with the depth to the 450 °C isotherm based on the same cooling plate model which successfully explains the heat flow, geoid anomaly, and subsidence of the oceanic lithosphere^{10,24}. The solid line in Fig. 1 shows the rigidities predicted by the depth to the 450 °C isotherm for this model with a 125-km plate thickness. For plate ages < 200 Myr at the time of loading, the oceanic and continental data are coincident and follow the 450 °C isotherm. The predicted trend, however, significantly departs from the continental data after 200 Myr. The continental data becomes compatible with the cooling plate model if the continental plate thickness is increased from 125 km to 200–250 km (dashed line, Fig. 1).

A thicker continental lithosphere relative to oceanic lithosphere has been suggested by many workers on the basis of continental heat flow data^{25–28}, thermometry of xenoliths and kimberlites^{24,30}, and shear wave propagation velocities^{31,32}. The thermal decay constant (τ) associated with the subsidence of sedimentary basins is proportional to the square of the lithospheric thickness. It has usually been assumed in models of basin subsidence^{1,33–38}, however, that the oceanic and continental thermal decay constants are similar. A continental lithospheric thickness of ~ 225 km implies a $\tau \sim 200$ Myr and a corresponding decrease in the amount of syn-rift subsidence of thermally subsiding basins. However, the best-fitting exponential to the tectonic subsidence will be a function of lateral heat flow, basin width, and the duration of rifting, and may be significantly less than τ . We suggest that the periods of apparent reactivation of cratonic basin subsidence (100–200 Myr, $\tau \sim 50$ Myr) may simply represent a punctuated history of a single episode of thermal subsidence (400–600 Myr, $\tau \sim 200$ Myr) rather than the superposition of many separate phases. The life span of any one phase is more a function of the history of plate reconfigurations rather than the thermal properties of the lithosphere.

The results of oceanic flexure studies are compatible with inferences from experimental rock mechanics^{34,39}. Goetze and Evans³⁹ developed a yield-stress envelope based on the rheology of olivine in which the yield strength in the upper part of the envelope was limited by brittle failure (Byerlee's law) while stress relaxation in the lower part results from thermally activated creep processes⁴⁰. This yield-stress model successfully predicts the flexural response for oceanic lithosphere of

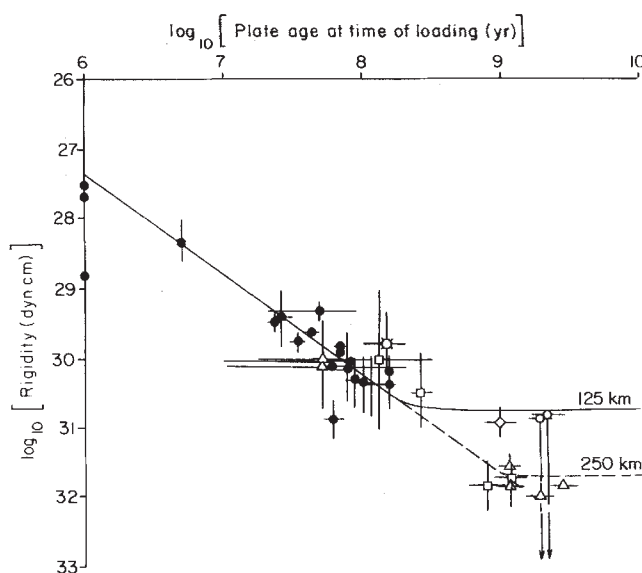
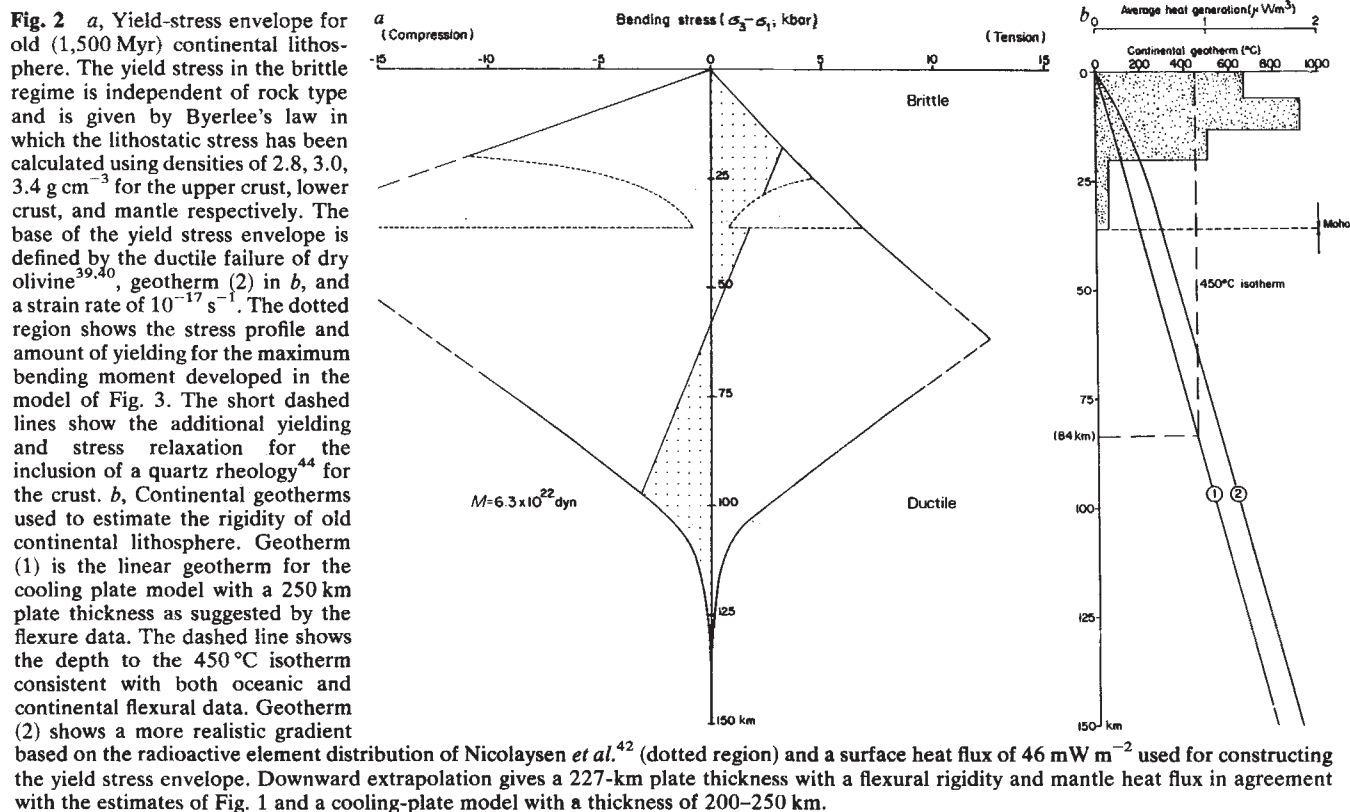


Fig. 1 A log-log plot of plate age at the time of loading (both oceanic and continental) versus the effective elastic rigidity of the lithosphere. Solid symbols denote oceanic rigidity values while open symbols refer to continental values (diamond, Michigan Basin; squares, foreland basins; triangles, ice loads; circles, minimization of gravity anomalies). The data were taken from ref. 1. and continental values determined by Karner and Watts¹² have also been included (Molasse and Appalachian foreland basins). Superimposed are the 450 °C isotherm for the cooling plate model¹⁰ with 125-km (solid line) and 250-km (dashed line) plate thicknesses. The apparent consistency between oceanic and continental rigidities suggest that the mechanical properties of both lithospheres are similar and that lithospheric rigidity increases with plate age. This tendency is inconsistent with predictions by the Maxwell visco-elastic model.

different age while predicting reasonable values for the stresses developed during loading. The deflections predicted by the yield-stress model closely agrees with those predicted by an elastic model in which the elastic thickness is defined by the depth to the 450 °C isotherm.

Following the numerical technique of Bodine *et al.*⁴¹, we test the appropriateness of the yield-stress model for the continents. To model realistically continental geotherms, we have included a radioactive element distribution (Fig. 2b) based on the study of Nicolaysen *et al.*⁴². This study sampled the Archaean granites of the updomed and overturned strata of the Vredefort structure, South Africa, to a presumed crustal depth of 20 km. The mantle heat flux of ~ 17 mW m⁻² obtained by Nicolaysen *et al.*⁴² for the Kaapvaal craton is in agreement with the geotherm of Fig. 2b and is consistent with a lithospheric thickness of 200–250 km. A low mantle heat flux and high temperature gradient within the upper crust also characterize results from the Precambrian shield of the Russian Kola Peninsula based on magneto-telluric and deep (~ 10 km) borehole measurements⁴³. Assuming that the South African geotherm is representative of the temperature structure of old continental lithosphere, we have constructed the yield stress envelope shown in Fig. 2a using a dry olivine rheology.

Figure 3 shows the resulting surface deflection and amount of brittle and ductile plastic failure within a cantilever plate for bending appropriate to flexural loading of the Indian foreland basement (Ganges Basin) by the Himalayan mountains. The effective elastic thickness of the plate varies from 116 km (where no yielding has occurred) to 88 km (at the point of maximum plastic failure). The dotted region in Fig. 2a shows the resulting stress distribution for the position of maximum bending moment (and hence curvature). As with oceanic flexure⁴¹, the large applied bending moments develop acceptable stress levels within the lithosphere (~ 3 kbar). The dashed line in Fig. 2a shows the modification of the yield stress envelope by the inclusion of a quartz rheology⁴⁴ for the crust. Since it is unlikely



that the entire crust is silicic⁴⁴, the model probably underestimates the strength of the crust. Figure 2*a* shows that, in any case, the quartz rheology produces only minor changes in the stress and bending moment developed within the lithosphere. The dotted region in Fig. 3*b* shows the approximate, restricted, region of ductile failure within the plate due to the quartz rheology and a significant (> 60 km) thick elastic layer which still remains below this failed region. It is apparent, therefore, that mantle rheology may be the prime control on continental flexure and that it is the olivine rheology which sustains the bending moment and elastic strains. The deflection resulting

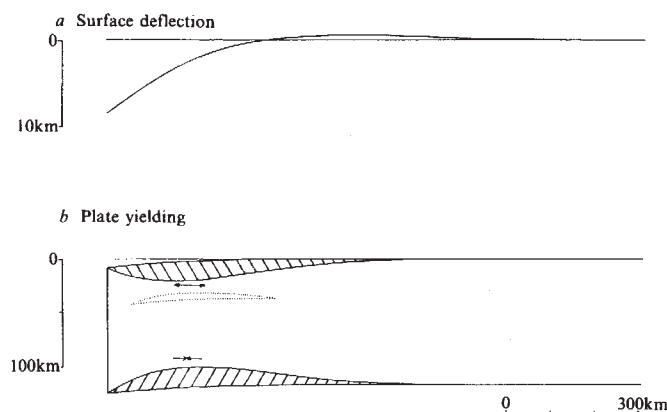


Fig. 3 Surface deflection (*a*) and yielding (*b*) within the mechanical thickness of the lithosphere for bending of a plate associated with an end load. The ~8,500 m deflection (Ganges Basin) corresponds to the bending of the ~1,500-Myr old Indian foreland due to the Himalayan mountain load. Upper and lower hatched regions show the degree of yielding within the brittle and ductile regimes of the plate. The dotted area outlines the additional yielding which would occur if a quartz rheology was applicable for the entire crust. Note that because the region of quartz flow is restricted, it will have a limited effect on plate strength. The surface deflection is consistent with a high rigidity ($\sim 10^{32}$ dyn cm, $T_e \sim 90$ –100 km) elastic plate.

from this modelling can be approximated by an elastic plate with an effective rigidity of $\sim 10^{32}$ dyn cm ($T_e \sim 100$ km). Thus, the large effective rigidity of old continental lithosphere can be modelled as the simple response to loading of a yield stress envelope based on dry olivine rheology.

Elastic stresses associated with the observed large rigidities result mainly from the storage of recoverable strains within the ductile region (Fig. 2*a*). Other geological results also suggest that significant recoverable strain can occur at substantial depths within the lithosphere and can be maintained over geological time. For example: (1) deep seismic sounding (DSS) in shield areas of low topographic relief (Baltic, Ukraine, and Indian) shows a rugged Moho topography (dominated by crustal faults) with approximate amplitudes of 20–30 km at mean depths of 40–50 km (refs 45, 46). (2) The recognition of low-angle normal faults characterizing extension in the Basin and Range province⁴⁷, COCORP results for the Wind River thrust⁴⁸, and the deep structure of the Moine and Flannen thrust zones^{49,50} all suggest that the entire continental crust (30–50 km) has been involved in faulting and thrusting. (3) Amorphous zones of earthquakes (epicentres 0–80 km), apparently unrelated to Benioff zones, characterize the Tibetan Plateau^{51–53}, Zagros mountains⁵⁴, Peruvian and Chilean Andes in the over-riding plate⁵⁵ and the Spanish Betic Cordillera⁵⁶. (4) Several deep intraplate earthquakes (30–50 km) occur within the ductile mantle of oceanic lithosphere⁵⁷. (5) Free-air gravity anomalies associated with Archaean and Proterozoic terrains⁵⁸ and young and old mountain ranges^{12,59,60} suggest that flexurally supported features are maintained over geological time. From these observations it would appear that creep processes have been ineffective in relaxing deviatoric stresses within the continental lithosphere at depths of ~20–100 km (Fig. 2).

We conclude that a Maxwell visco-elastic model, which predicts a decrease in lithospheric rigidity with time, is inappropriate for either continental or oceanic lithospheres. However, because of the invariance of rigidity once acquired, the deflection and hence free-air gravity anomaly which results from loading are adequately approximated by an elastic model. The actual mechanical properties of the lithosphere can be

expected to be considerably more complex than suggested by this model, but, as discussed by Bodine *et al.*⁴¹, and Goetze and Evans³⁹, and extended to the continental lithosphere in this study, the flexural response of a yield-stress model for the lithosphere is consistent with the elastic model while predicting acceptable values for the stresses developed during loading. In addition, it appears that the long-term thermal behaviour of the continental lithosphere is governed by a cooling-plate model with a 200–250 km plate thickness.

We thank K. Jacobs, N. Ribe, J. Weissel, J. Cochran, A. Watts, and K. Hegarty for reading the manuscript and W. Haxby and J. Weissel for stimulating discussions. This work was supported by an Australian Public Service Postgraduate Scholarship (G.D.K.), NATO Postdoctoral fellowship and Lamont-Doherty Geological Observatory Institutional Funds (M.S.S.), and NSF contract EAR 81-09473. L-DGO contribution 3492.

Received 24 November 1982; accepted 1 June 1983.

1. Watts, A. B., Karner, G. D. & Steckler, M. S. *Phil. Trans. R. Soc. A* **305**, 249–281 (1982).
2. Beaumont, C. *Geophys. J. R. astr. Soc.* **65**, 291–329 (1981).
3. Watts, A. B. *Nature* **297**, 460–474 (1982).
4. Beaumont, C. *Geophys. J. R. astr. Soc.* **55**, 471–498 (1978).
5. Sleep, N. H. *Geophys. J. R. astr. Soc.* **24**, 352–350 (1971).
6. Watts, A. B. *J. geophys. Res.* **83**, 5989–6004 (1978).
7. Watts, A. B., Bodine, J. H. & Steckler, M. S. *J. geophys. Res.* **85**, 6369–6376 (1980).
8. Cazenave, A., Lago, B., Dominti, K. & Lambeck, K. *Geophys. J. R. astr. Soc.* **63**, 233–354 (1980).
9. Caldwell, J. G. & Turcotte, D. L. *J. geophys. Res.* **84**, 7572–7576 (1979).
10. Parsons, B. & Sclater, J. G. *J. geophys. Res.* **82**, 803–827 (1977).
11. Walcott, R. I. *J. geophys. Res.* **75**, 3941–3954 (1970).
12. Karner, G. D. & Watts, A. B. *J. geophys. Res.* (in the press).
13. Sleep, N. H. & Sloss, L. L. *J. geophys. Res.* **83**, 5815–5819 (1978).
14. McCallister, R. H., Boctor, N. Z. & Hinz, W. J. *J. geophys. Res.* **83**, 5825–5831 (1978).
15. Van der Voo, R. & Watts, D. R. *J. geophys. Res.* **83**, 5844–5848 (1978).
16. Lewis, B. T. R. & Dorman, L. M. *J. geophys. Res.* **75**, 3367–3386 (1970).
17. McNutt, M. K. & Parker, R. L. *Science* **199**, 773–775 (1978).
18. McNutt, M. K. *J. geophys. Res.* **85**, B11, 6377–6396 (1980).
19. Stephanson, R. & Beaumont, C. in *Mechanism of Continental Drift and Plate Tectonics* (eds Davies, G. F. & Runcorn, R. K.) 111–122, (Academic, New York, 1980).
20. Forsyth, D. W. *Rev. Geophys. Space Phys.* **17**, 1109–1114 (1979).
21. Forsyth, D. W. *EOS* **62**, 1032 (1981).
22. Haxby, W. F., Turcotte, D. L. & Bird, J. M. *Tectonophysics* **36**, 57–75 (1976).
23. Molnar, P. & Tapponnier, P. *Earth planet. Sci. Lett.* **52**, 107–114 (1981).
24. Sandwell, D. T. & Schubert, G. *J. geophys. Res.* **87**, 3949–3958 (1982).
25. Oxburgh, E. R. *Phil. Trans. R. Soc. A* **301**, 337–346 (1981).
26. Davies, G. F. & Strebeck, J. W. *Geophys. J. R. astr. Soc.* **69**, 623–634 (1982).
27. Pollack, H. N. A. *Rev. Earth planet. Sci.* **10**, 459–481 (1982).
28. Vitarello, I. & Pollack, H. N. *J. geophys. Res.* **85**, 983–995 (1980).
29. Gurney, J. J., Harte, B. & Cox, K. G. *Phys. Chem. Earth* **9**, 507–523 (1977).
30. Eggler, D. H., McCallum, M. E. & Smith, C. B. in *Proc. 2nd int. Kimberlite Conf.*, Santa Fe (American, Washington, 1977).
31. Toksoz, M. N. & Anderson, D. L. *J. geophys. Res.* **71**, 1649–1658 (1966).
32. Jordan, T. H. *Rev. Geophys. Space Phys.* **13**, 1–12 (1975).
33. Beaumont, C., Keen, C. E. & Boutilier, R. *Geophys. J. R. astr. Soc.* **70**, 667–715 (1982).
34. Brunet, M. F. & Le Pichon, X. *J. geophys. Res.* **87**, 8547–8560 (1982).
35. Dewey, J. F. & Pittman, W. C. *Phil. Trans. R. Soc. A* **305**, 145–148 (1982).
36. Turcotte, D. in *Geodynamics Series Vol. 1* (eds Bally, A. W., Bender, P. L., McGetchin, T. R. & Walcott, R. I.) (AGU, Washington DC, 1980).
37. Sclater, J. G. & Christie, P. A. *J. geophys. Res.* **85**, 3711–3739 (1980).
38. McKenzie, D. *Earth planet. Sci. Lett.* **40**, 25–32 (1978).
39. Goetze, C. & Evans, B. *Geophys. J. R. astr. Soc.* **59**, 463–478 (1979).
40. Goetze, C. *Phil. Trans. R. Soc. A* **288**, 99–119 (1978).
41. Bodine, J. H., Steckler, M. S. & Watts, A. B. *J. geophys. Res.* **86**, 3695–3707 (1981).
42. Nicolaysen, L. O., Hart, R. J. & Gale, N. H. *J. geophys. Res.* **86**, 10653–10661 (1981).
43. Zhamaletdinov, A. A. in *Deep Electromagnetic Sounding using the MGD Impulse Generator* (in Russian) (ed. Kirova, S. M.) 35–46 (The Geological Institute (Kola Branch) of the Soviet Academy of Sciences, Apatity, 1982).
44. Brace, W. F. & Kohlstedt, D. L. *J. geophys. Res.* **85**, 6248–6252 (1980).
45. Sollogub, V. B. *et al. Tectonophysics* **20**, 1–33 (1973).
46. Chowdhury, K. R. & Hargraves, R. B. *Nature* **291**, 648–650 (1981).
47. Wernke, B. *Nature* **291**, 645–648 (1981).
48. Smithson, S. B., Brewer, J., Kaufman, S. & Oliver, J. *Geology* **6**, 648–652 (1978).
49. Soper, N. J. & Barber, A. J. *J. geol. Soc. Lond.* **139**, 127–138 (1982).
50. Smythe, D. K. *et al. Nature* **299**, 338–341 (1982).
51. Chen, W.-P. & Molnar, P. *J. geophys. Res.* **82**, 2945–2969 (1977).
52. Gupta, H. K. & Bhatia, S. C. in *Geodynamics Series Vol. 3* (eds Gupta, H. K. & Delaney, F. M.) (AGU, Washington DC, 1981).
53. Ni, J. & Barazangi, M. *Geophys. J. R. astr. Soc.* **72**, 665–689 (1983).
54. Bird, P. *Tectonophysics* **50**, 307–336 (1978b).
55. Isacks, B. L. & Barazangi, M. in *Island Arcs, Deep Sea Trenches and Back-arc Basins* (Maurice Ewing Ser. 1, AGU, Washington DC, 1977).
56. Banda, E. & Vegas, R. *Pap. at. Workshop Evolution of Extensional Basins within Regions of Compression, with Emphasis on the Intra-carpethians, Vesztem, Hungary* (1982).
57. Wiens, D. A. & Stein, S. *J. geophys. Res.* (submitted).
58. Fountain, D. M. & Salisbury, M. H. *Earth planet. Sci. Lett.* **56**, 263–277 (1981).
59. Kahle, H.-G., Klingele, E., Mueller, St. & Egloff, R. *Pure Appl. Geophys.* **114**, 479–494 (1976).
60. Diment, W. H., Urban, T. C. & Revetta, R. A. in *The Nature of the Solid Earth* (eds Robertson, E. C. *et al.*) (McGraw-Hill, New York, 1972).

Lithosphere–asthenosphere decoupling at spreading centres and initiation of obduction

John G. Spray

Department of Earth Sciences, University of Cambridge, Cambridge CB2 3EQ, UK

One of the outstanding problems for geologists studying ophiolites is to understand how obduction¹ takes place. Current opinion interprets ophiolites as fragments of oceanic lithosphere, and obduction as a series of events which result in ophiolites being tectonically transported from an oceanic to a continental location. It is the early stages of the obduction process which remain the most obscure; in particular, how ophiolites become detached from their oceanic site of generation. Any attempt at understanding obduction requires the identification of key features which may indicate provenance and likely decoupling mechanisms. Six such features are considered here: ophiolite dimensions, internal structure, chemistry, petrology, the presence of metamorphic soles and their radiometric ages relative to those of ophiolite crystallization. Taken together the evidence suggests that certain ophiolites originate by the overthrusting of hot, thin lithosphere in the vicinity of spreading centres by decoupling along the elevated lithosphere–asthenosphere boundary.

Ophiolites characteristically form linear belts extending for up to several hundred kilometres along plate sutures. In contrast, their width rarely exceeds 100 km and their thickness is usually <15 km. These approximate dimensions remain generalizations because many ophiolites have undergone deformation and metamorphism during and after continental emplacement which has resulted in their tectonic thinning or thickening. Nevertheless, the better preserved examples, such as the Semail Complex (Oman) and Bay of Islands Complex (Newfoundland), do possess tabular dimensions where their length > width > thickness. Significantly, even allowing for the effects of tectonism, no ophiolite complex has yet been found thicker than 15 km, indicating that decoupling occurs within the uppermost mantle at a level equal to or less than this depth. We should therefore be aware of the possible existence of a potential detachment level and aim to identify it.

The internal structure of ophiolites has been likened to that of ocean crust and upper mantle because of the similarities shown in their respective seismic profiles and sampled lithologies (Fig. 1). However, there are differences which indicate that many ophiolites are not formed at normal mid-ocean ridges. The supra-Moho sequence of some ophiolites is apparently thinner than that of mature oceanic crust and the individual units within them may also show variable thicknesses: sheeted dykes and an overlying pelagic veneer may be absent or poorly developed and gabbros of layer 3 thinned, whereas lavas of the extrusive unit are occasionally found thickened². Variable crustal thicknesses have been explained in several ways, including variation in spreading rate (for example, the production of thin sheeted dyke but thick gabbro units have been predicted during fast spreading³), or the ophiolite being derived from the spreading centre itself, where crustal units may occur incompletely developed⁴.

While there clearly remain problems in precisely comparing the structure of ophiolites with that of oceanic crust, mainly because of the inaccessibility of the latter, indications that ophiolites may be anomalous crust also emerge from geochemical studies. Some ophiolites possess a chemical signature typical of mid-ocean ridge basalts (MORB, such as the Bay of Islands Complex⁵) but many show evidence of calc-alkaline tendencies (Troodos⁶, Oman⁷, Pindos⁸). Along with

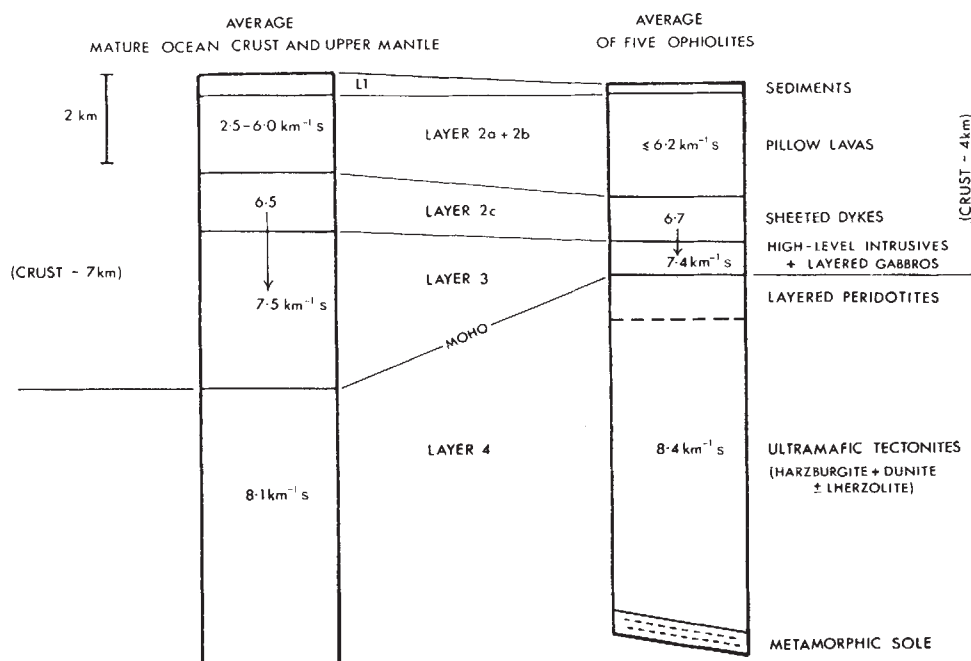


Fig. 1 Tentative seismic and layer thickness comparisons between mature ocean crust with upper mantle and model ophiolite (mean thicknesses determined using the Bay of Islands, Papua, Semail, Troodos and Vourinos ophiolites⁴). Approximate P-wave velocities are indicated for each layer³¹. The greater thickness of Layer 2(a + b) in the model ophiolite may be due to tectonic thickening. The high-level intrusives may include gabbro, diorite, tonalite and trondjhaemite lithologies.

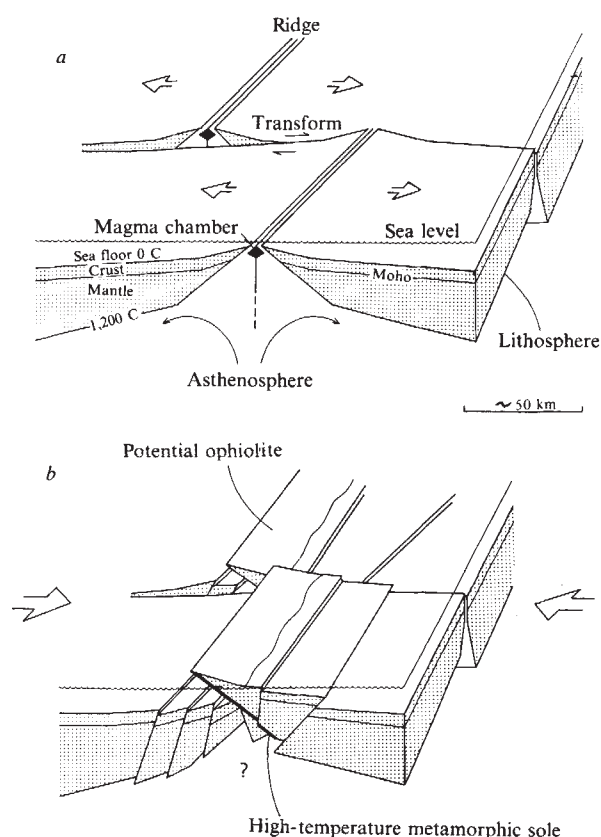


Fig. 2 *a*, Schematic section of an oceanic spreading centre. During extension new lithosphere is formed at the ridge which thickens with age by progressive underplating. The rigid lithosphere is bounded by the 0 and 1,200°C isotherms and overlies the plastic asthenosphere. *b*, The extensional regime is replaced by one of compression. Overthrusting takes place at the ridge along the shallow lithosphere-asthenosphere boundary along which a high-temperature metamorphic sole is formed. Subsequent compressive movements may result in the initiation of subduction, either at the ridge or within older lithosphere at the basin margin. Continuing subduction may result in the overthrust wedge being transported onto a continental margin as an ophiolite.

the presence of high-level tonalites and trondjhaemites, the calc-alkaline association is substantiated by the common but often poorly documented occurrence of granitoid intrusions within ophiolites (for example, at Guevgueli, Greece⁹ and in the southern Andes¹⁰). These factors have led many workers to favour a supra-subduction zone site of origin for ophiolites, involving fore-arc¹¹, arc¹², back-arc¹³ and 'arc-less' marginal basin¹⁴ variants.

A further constraint on ophiolite provenance is provided by the type of peridotite which constitutes the mantle tectonites. Most tectonites appear dominated by depleted harzburgite and dunite, but many contain lherzolite, representing less depleted mantle, particularly towards ophiolite bases. Some mantle sequences even appear to comprise predominantly lherzolite¹⁵. High-pressure mafic segregations may also occur in association with the lherzolites as variably deformed veins, dykes and bands¹⁶. These have been interpreted as trapped primitive melts which escaped significant fractionation due to their rapid ascent, coupled with a failure to access crustal magma reservoirs¹⁶. Presence of the mafic layer-lherzolite association is considered to be the result of lower degrees of partial melting than is normally associated with mature oceanic spreading centres, beneath which the efficient extraction of basaltic melt and concomitant depletion of the mantle take place. This has led several workers to favour a marginal continental site of origin for lherzolitic ophiolites, their generation being coincident with the inception of rifting and the early stages of basin evolution or the development of small oceanic basins^{16,17}.

With regard to the decoupling process, the base of many ophiolites is marked by a thin zone (usually <500 km thick) exhibiting the effects of dynamothermal metamorphism (Fig. 1). It has been shown that these metamorphic soles are acquired during the early stages of ophiolite movement and their study has consequently provided important insights into decoupling conditions¹⁸. Complete soles may show evidence of granulite facies metamorphism and even anatexis¹⁸. Such high temperatures can only be realized by the displacement of a hot body, even if additional heat is supplied by friction during thrusting¹⁹. This is substantiated by study of the age relations between igneous crystallization and metamorphic sole formation. Ages have been well-documented for the Newfoundland⁵, Greek-Yugoslavian²⁰ and Oman²¹ ophiolites. After necessary corrections have been made using Steiger and Jager age constants²², it is apparent that the timing of igneous crystallization

and metamorphic sole formation of these ophiolites is virtually synchronous, or they occur within ~5 Myr of each other¹⁸. Significantly, this limits the location of their initial decoupling and overthrusting to the proximity of their respective spreading centres. Whether or not other ophiolite complexes show similar age relations must await further geochronological investigations, in particular the determination of igneous crystallization ages which are presently poorly constrained.

The above evidence indicates that: (1) those ophiolites possessing thin crustal sequences and/or calc-alkaline associations are indicative of a supra-subduction zone provenance; (2) the presence of significant amounts of lherzolite within ophiolite mantle sequences suggests derivation during the early stages of rifting or within a small oceanic basin. Points (1) and (2) are not mutually exclusive, in fact both are in support of such ophiolites originating in some form of 'marginal' basin rather than at a mature mid-ocean ridge. Verification of these inferences must await further detailed chemical and petrological studies and mapping of ophiolites and oceanic lithosphere. Notwithstanding this, some of the most significant advances in our understanding of how and from where ophiolites originate have arisen from geochronological investigations: in the few cases where both igneous crystallization and metamorphic sole ages are known they are found to be broadly synchronous. This crucial evidence may provide the key to identifying the ophiolite detachment level.

It is proposed that the most likely decoupling horizon is the lithosphere–asthenosphere boundary beneath the spreading centre. This boundary defines the transition from rigid to plastic behaviour within the upper mantle, broadly corresponding to a low velocity zone for both S and P waves (Fig. 2a). The boundary approximates the 1,200 °C isotherm which lies at ~100 km depth for mature lithosphere. However, the boundary is uniquely elevated beneath spreading centres and may thus facilitate the tectonic detachment of a hot, immature lithospheric wedge on ridge compression (Fig. 2b). The amount of overthrusting is likely to be controlled by the compressive force and by lithospheric thickness which is itself governed by the rate of underplating. This has been estimated by several workers²³: simple boundary layer cooling models give an (age)^{1/2} law for lithospheric thickness if younger than ~60 Myr (refs 24, 25). For example, thicknesses of up to 25 km are expected for lithosphere <5-Myr-old, though this is clearly dependent on the spreading rate. At a half rate of 1.5 cm yr⁻¹, spreading for 5 Myr would yield a potential maximum overthrust ophiolite width of ~75 km, a width comparable with that of many ophiolite belts.

These conclusions agree with those of previous workers including Christensen and Salisbury⁴ (who favoured ridge-trench collision as a method of procuring ophiolites), Boudier and Coleman²¹ and Boudier *et al.*²⁶. In addition, overthrusting lithosphere along the lower limit of its elastic strength (~600 °C isotherm) has been proposed and remains a viable alternative site of origin^{26,27}. Note that ridge–continent collision need not be a prerequisite to overthrusting: there is increasing evidence that many ophiolites are initially displaced within the oceanic realm, in some cases tens of megayears before their final continental emplacement²⁸. For these reasons it is to the structure of spreading centres that our attentions should be focused, the full complexities of which remain to be revealed (see, for example, ref. 29): young, thin lithosphere and the proximity of fracture zones³⁰ clearly provide an ideal decoupling location for the initiation of obduction.

This work was completed during tenure of an Earth Sciences Research Fellowship from King's College Cambridge which is gratefully acknowledged. I thank I. G. Gass for critically reviewing the manuscript. Cambridge Earth Sciences contribution 288.

Received 21 February; accepted 31 May 1983.

1. Coleman, R. G. *J. geophys. Res.* **76**, 1212–1222 (1971).
2. Alabaster, T. *et al. Ophiolite Symp.*, Cyprus, 751–757 (1980).
3. Sleep, N. H. & Rosendahl, B. R. *J. geophys. Res.* **84**, 6831–6839 (1979).
4. Christensen, N. I. & Salisbury, M. H. *Rev. geophys. Space Phys.* **13**, 57–86 (1975).
5. Jacobsen, S. B. & Wasserburg, G. J. *J. geophys. Res.* **84**, 7429–7445 (1979).

6. Pearce, J. A. *Ophiolite Symp.*, Cyprus, 261–272 (1980).
7. Pearce, J. A. *et al. Phil. Trans. R. Soc. A* **300**, 299–317 (1981).
8. Capedri, S. *et al. Contr. Miner. Petrol.* **74**, 189–200 (1980).
9. Bebie, J. *Nature* **270**, 232–234 (1977).
10. Dalziel, W. D., de Wit, M. J. & Palmer, K. F. *Nature* **250**, 291–294 (1974).
11. Karig, D. E. *Spec. Publ. geol. Soc. Lond.* **10**, 563–576 (1982).
12. Miyashiro, A. *Earth planet. Sci. Lett.* **19**, 218–224 (1973).
13. Hawkins, J. W. *Ophiolite Symp.*, Cyprus, 244–254 (1980).
14. Smith, A. G. & Spray, J. G. *Spec. Publ. geol. Soc. Lond.* (in the press).
15. Nicolas, A. *et al. Nature* **294**, 414–417 (1981).
16. Spray, J. G. *Nature* **299**, 524–528 (1982).
17. Menzies, M. A. & Allen, C. *Contr. Miner. Petrol.* **45**, 197–213 (1974).
18. Spray, J. G. *Spec. Publ. geol. Soc. Lond.* (in the press).
19. Graham, C. M. & England, P. C. *Earth planet. Sci. Lett.* **31**, 142–152 (1976).
20. Spray, J. G. *et al. Spec. Publ. geol. Soc. Lond.* (in the press).
21. Boudier, F. & Coleman, R. G. *J. geophys. Res.* **86**, 2573–2592 (1981).
22. Steiger, R. H. & Jager, E. *Earth planet. Sci. Lett.* **36**, 359–362 (1977).
23. Forsyth, D. W. *Tectonophysics* **38**, 89–118 (1977).
24. Parker, R. L. & Oldenburg, D. W. *Nature* **242**, 137–139 (1973).
25. Parsons, B. & Sclater, J. G. *J. geophys. Res.* **82**, 803–827 (1977).
26. Boudier, F., Nicolas, A. & Bouchez, J.-L. *Nature* **296**, 825–828 (1982).
27. Nicolas, A. & Le Pichon, X. *Earth planet. Sci. Lett.* **46**, 397–406 (1980).
28. Spray, J. G. & Roddick, J. C. *Contr. Miner. Petrol.* **72**, 43–55 (1980).
29. MacDonald, K. C. & Fox, P. J. *Nature* **302**, 55–58 (1983).
30. Fox, P. J., Detrick, R. S. & Purdy, G. M. *Ophiolite Symp.*, Cyprus, 161–168 (1980).
31. Salisbury, M. H. & Christensen, N. I. *J. geophys. Res.* **83**, 805–817 (1978).

Synthesis of Na- and K-clinoptilolite endmembers

Cheng-Hang Chi & L. B. Sand

Department of Chemical Engineering, Worcester Polytechnic Institute, Worcester, Massachusetts 01609, USA

We report here the first synthesis of endmembers Na- and K-clinoptilolite. Na, K-clinoptilolites, formed principally in nature by the alteration of siliceous volcanic ash and tuffs deposited in the ocean or in continental saline, alkaline lakes, are the most abundant zeolites occurring as minerals. However, it has been found difficult to delineate in the laboratory the conditions which lead to the crystallization of these minerals at low-temperature and low-pressure conditions. In the reactant systems studied and in a temperature range 120–200 °C at autogenous pressure, Na-clinoptilolite crystallized with the addition of a small amount (1%) of seed crystals, and K-clinoptilolite was synthesized with or without seed. Intermediate Na, K-clinoptilolites also crystallized in the seeded and unseeded systems. Coexisting phases, depending on the reactant composition, included mordenite, phillipsite, gismondine (zeolite P₁) and K-feldspar. An experimental basis has now been established to determine the kinetics and paragenesis of this zeolite occurring in very large amounts in ocean and continental sediments.

In spite of its great abundance in nature, clinoptilolite has been found difficult to synthesize in the laboratory. The two principal contributions have been by Hawkins *et al.*¹ who synthesized Na,K-clinoptilolite from volcanic glass at 130–150 °C under the high pressure of 1 kbar and by Goto² who also synthesized Na,K-clinoptilolite but from chemical reactants at 200 °C and autogenous pressure (~15 atm). Alkaline earth heulandite–clinoptilolites, which occur in metamorphic and hydrothermal deposits, had been synthesized earlier^{3,4} at high temperature and pressure. None of the previous work has reported the kinetics of crystallization of clinoptilolite.

A typical reactant molar composition of Na-clinoptilolite and K-clinoptilolite is 2.1MOH–Al(OH)₃–5SiO₂–52.5H₂O, where M is Na or K. Reactants used in our laboratory synthesis were reagent-grade sodium (Fisher) or potassium (J. T. Baker) hydroxide, dried aluminium hydroxide gel (Reheis F-2000), aqueous colloidal silica (Du Pont Ludox-HS) and deionized water. Syntheses were carried out in small (10 ml) stainless steel autoclaves with Teflon liners at 120–200 °C and autogenous pressure. The addition of seed crystals (1–10% Sheaville, Oregon clinoptilolite) was found to be important in the Na system. Although K-clinoptilolite and Na,K-clinoptilolite were synthesized without seeding, a small amount (1%) of seed crystals improved both the degree and the rate of crystallization.

Table 1 Typical syntheses of clinoptilolite with coexisting phases

Run no.	Reactant composition NaOH/KOH/Al(OH) ₃ /SiO ₂ /H ₂ O	Seeds (wt%)	Temperature (°C)	Time (h)	Results
225	2.1:0:1:5:52.5	10	120	300	Na-clinoptilolite (100%)
227	2.1:0:1:5:52.5	10	140	64	Clinoptilolite (90%) Mordenite (10%)
218	1.68:0:1:4:42	10	175	15	Clinoptilolite (50%) Gismondine (20%) Mordenite (5%)
207	0.53:1.57:1:5:52.5	0	175	76	Clinoptilolite (75%) Mordenite (10%) Phillipsite (5%)
211	0:2.1:1:5:52.5	1	175	94	Clinoptilolite (95%) K-feldspar (5%)
224	0:2.1:1:5:52.5	10	195	37	K-clinoptilolite (100%)

The kinetics of crystallization of both Na-clinoptilolite and K-clinoptilolite were studied and the results are shown in Fig. 1. At the same temperatures and pressures, the kinetics of crystallization of K-clinoptilolite was found to be two to four times slower than that of Na-clinoptilolite. The rates of crystallization, defined as the rate of conversion at 50% of the total conversion level⁵, were obtained at different temperatures. The dependence of the rates on temperature were expressed by the Arrhenius equation to estimate the apparent activation energies for crystallization of Na- and K-clinoptilolites. The values obtained are 13.8 for Na-clinoptilolite and 14.5 kcal gmol⁻¹ for K-clinoptilolite. Note that these values are close to that of synthetic Na-mordenite⁵ (15 kcal gmol⁻¹). Mordenite is a mineral which commonly occurs with clinoptilolite in nature.

Depending on the reactant composition, mordenite, phillipsite, gismondine (zeolite P₁) and K-feldspar were found coexisting with clinoptilolite (see Table 1). Work is in progress to delineate the synthesis conditions in more detail, including the rate of nucleation and the metastable behaviour of clinoptilolite. These results would provide an experimental basis for determining the kinetics and paragenesis of this most abundant of zeolite minerals in nature.

The synthesized samples were analysed by X-ray powder diffraction with the addition of 1% aluminium powder as internal standard. Table 2 summarizes the *d*-spacings and intensities for both Na- and K-clinoptilolites.

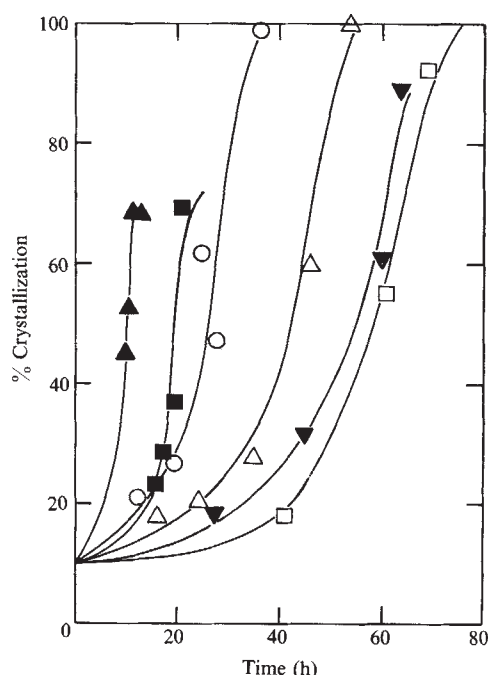


Fig. 1 Crystallization curves of Na-clinoptilolite (shaded symbols) and K-clinoptilolite (open symbols) in seeded systems. ○, 195 °C; △, 175 °C; □, 160 °C; ▽, 140 °C.

The scanning electron photomicrograph of K-clinoptilolite synthesized in a seeded system shows the platy morphology of the crystals, as shown in Fig. 2. The crystal sizes are typically 1–5 μm in a seeded system. Na-clinoptilolite synthesized in the seeded system has a similar morphology. Figure 3 shows the Na, K-clinoptilolite synthesized in an unseeded system. The crystal sizes are typically 10–15 μm.

We conclude that the surprising difficulty in synthesizing the alkali clinoptilolites was due to the limited crystallization conditions found experimentally. One might assume that it would crystallize from a wide range of conditions in view of its abundant natural occurrences, but the apparent contradiction is attributed to the fact that the chemical composition of clinoptilolite is almost identical to that of the parent volcanic glass. The experimental data reveal that clinoptilolite is the low-temperature phase and mordenite with similar composition but higher activation energy is the higher-temperature phase.

The crystallization data shown in Fig. 1 are consistent with the observation of Utada⁶ that K-rich clinoptilolites occur

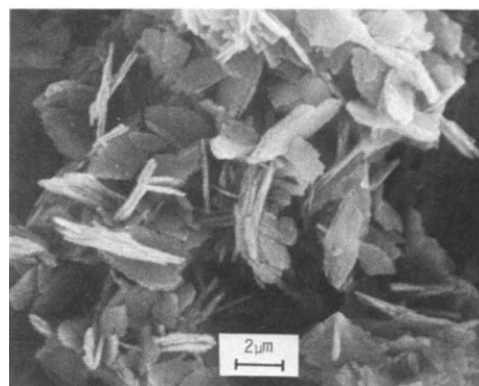


Fig. 2 Scanning electron photomicrograph of K-clinoptilolite synthesized in a seeded system.

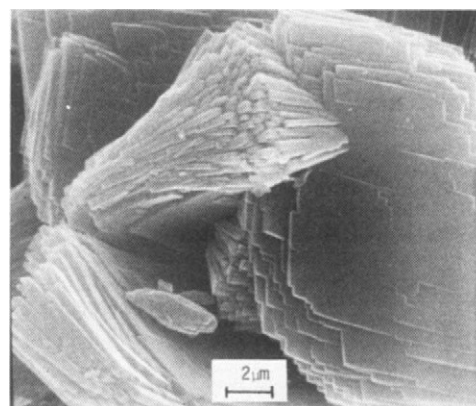


Fig. 3 Scanning electron photomicrograph of Na,K-clinoptilolite synthesized in an unseeded system.

Table 2 X-ray powder diffraction data

Na-clinoptilolite		K-clinoptilolite	
<i>d</i> (Å)	<i>I</i> / <i>I</i> ₀	<i>d</i> (Å)	<i>I</i> / <i>I</i> ₀
9.05	75	9.00	100
7.96	31	7.96	26
6.86	16	6.85	10
5.96	7	5.97	8
5.25	16	5.24	8
5.15	27	5.15	19
4.66	18	4.66	14
4.36	9	4.36	6
3.990	100	3.980	68
3.921	51	3.912	40
3.729	5	3.729	7
3.555	20	3.552	14
3.422	48	3.422	29
3.329	18	3.326	17
3.172	37	3.167	24
3.128	23	3.125	18
3.078	18	3.074	15
2.985	49	2.978	49
2.806	32	2.807	20
2.735	19	2.733	19
2.680	11	2.684	6
2.444	11	2.445	12

without mordenite, whereas Na-rich clinoptilolites occur with mordenite.

The contributions of Mr Glenn W. Dodwell on X-ray analysis, Dr Rosemarie Szostak on the SEM, and the financial support of NSF under grant no. CPE8017452 are gratefully acknowledged.

Received 1 February; accepted 4 May 1983.

1. Hawkins, D. B., Sheppard, R. A. & Gude, A. J. *3rd Natural Zeolites: Occurrence, Properties, Use* (eds Sand, L. B. & Mumpton F. A.) 337–343 (Pergamon, Oxford, 1978).
2. Goto, Y. *Am. Miner.* **62**, 330–332 (1977).
3. Koizumi, M. & Roy, R. J. *Geol.* **68**, 41–53 (1960).
4. Hawkins, D. B. *Mater. Res. Bull.* **2**, 951–958 (1967).
5. Culfaz, A. & Sand, L. B. *Adv. Chem. Ser.* **121**, 140–151 (1973).
6. Utada, M. *Scient. Pap. Coll. gen. Educ.*, Tokyo **20**, 191–262 (1970).

Exsolution of cummingtonite, actinolite and sodic amphibole in hornblende in high-pressure metamorphism

Y. Yamaguchi*, H. Shibakusa† & K. Tomita‡

* Department of Geology, Shimane University, Matsue 690, Japan

† Kobe Ikuei High School, Kobe 653, Japan

‡ Department of Geology and Mineralogy, Kyoto University, Kyoto 606, Japan

Exsolution phenomena in amphiboles indicate the existence of miscibility gaps^{1,2}. The phase relationships of hornblende, the commonest amphibole in igneous and metamorphic rocks, to the other monoclinic amphiboles such as cummingtonite, actinolite and sodic amphibole has particular petrological importance. In these systems a solvus has been directly confirmed by exsolution in the hornblende–cummingtonite system^{3–5}. Chemical discontinuities suggesting miscibility gaps have been shown for other systems^{6–18}. In the hornblende–actinolite system, which was recently confirmed to have a solvus by experimental synthesis¹⁹, fine lamellar texture that is too fine for microprobe resolution is expected to result from exsolution^{9,10}. However, there has been no report of amphibole exsolution from high-pressure metamorphic rocks. We have found exsolution lamellae of three kinds of clin amphibole, cummingtonite, actinolite and sodic amphibole, in a zoned hornblende from high-pressure metamorphic rock, and present here direct evidence to confirm the miscibility gaps between hornblende and these three clin amphiboles in high-pressure metamorphic conditions. The genesis of exsolution lamellae discussed below is based on back-scattered electron-scanning images and microprobe analyses.

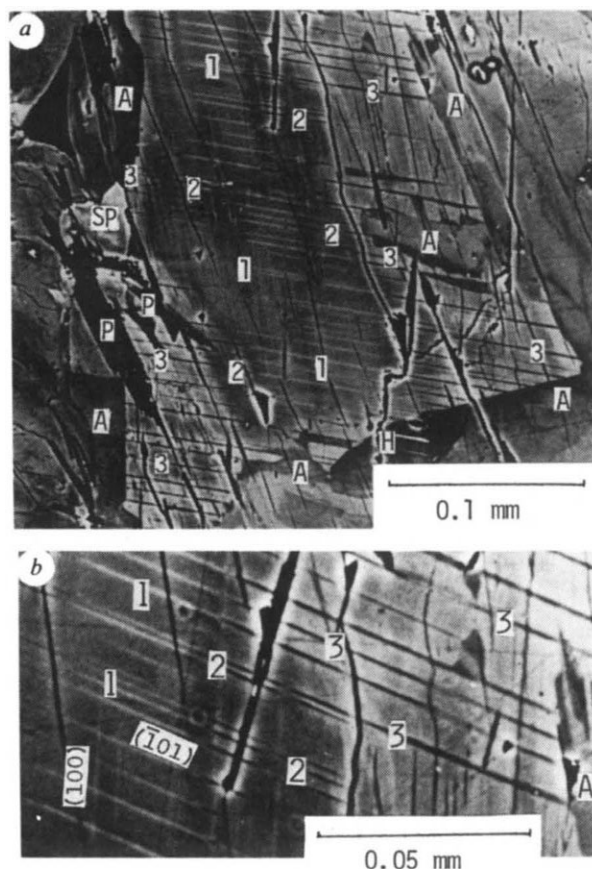


Fig. 1 *a*, Back-scattered electron scanning image of a zoned hornblende (by JXA-733 Microanalyser under 25 kV accelerating voltage and 0.02 μ A specimen current). 1, Cummingtonite lamellae (white lamellae in dark host); 2, sodic amphibole lamellae (dark lamellae in dark host); 3, actinolite lamellae (dark lamellae in grey to light host). H, hornblende lamellae very rarely occurring in rimming actinolite; A, actinolite; SP, omphacite; P, phengite (see text). *b*, Enlarged micrograph of the upper part in the same image as *a*. The extensions of cummingtonite lamellae changing into sodic amphibole lamellae and actinolite lamellae. Fine (100) cummingtonite lamellae occur in the area containing (101) cummingtonite lamellae, locally showing tweed texture.

The zoned hornblende has been found in a small lenticular epidote amphibolite (200 $\times$ 800 m) enclosed in glaucophane schist, exposed near Horokanai in the mélange zone of the Kamuikotan tectonic belt, Hokkaido, Japan. The epidote amphibolite underwent the same glaucophane schist facies metamorphism as that of the neighbouring glaucophane schist^{20–23}, with resultant high-pressure low-temperature mineral association comprising actinolite, phengite, omphacite, quartz and albite in its feldspathic band. The rock specimen containing the zoned hornblende mainly consists of minerals of the epidote amphibolite facies, comprising katophoritic hornblende, epidote, Mg-rich chlorite and rutile. The rock displays pronounced disequilibrium texture, due to overprinting of the glaucophane schist facies metamorphism. The katophoritic hornblende is generally rimmed by actinolite, omphacite ($\text{Di}_{44}\text{Jd}_{34}\text{Hd}_{16}\text{Ac}_6$) and phengite.

Figure 1 shows the back-scattered electron scanning images of the zoned hornblende. Fine clin amphibole lamellae are seen throughout the hornblende, with a maximum thickness of $\sim 2.0 \mu\text{m}$. The lamellae were optically determined to be parallel to (101). Microprobe analyses reveal the lamellae to be (1) cummingtonite in the core; (2) sodic amphibole in the outer core; and (3) actinolite in the rim. The actinolite lamellae very rarely extend into the rimming actinolite and change to hornblende lamellae, but the actinolite is mostly free from exsolution. The rim hornblende, containing the actinolite lamellae, has a chemically sharp boundary with the rimming actinolite.

Table 1 Structural formulae of clinoamphiboles, calculated from representative analyses

	A	Lamellae			Host			Lamellae + host		
		1L	2L	3L	1H	2H	3H	1B	2B	3B
Si	7.629	7.73	7.73	7.71	6.561	6.583	6.598	6.700	6.64	6.65
Al(IV)	0.371	0.27	0.27	0.29	1.439	1.417	1.402	1.300	1.36	1.35
Al(VI)	0.108	0.06	1.08	0.08	0.591	0.551	0.570	0.549	0.58	0.54
Ti	0.007	0.01	0.00	0.01	0.109	0.116	0.117	0.102	0.11	0.11
Fe	1.037	1.43	1.02	1.01	1.054	0.927	1.222	1.107	0.93	1.21
Mn	0.035	0.04	0.00	0.00	0.022	0.014	0.030	0.022	0.01	0.03
Mg	3.992	5.38	3.19	4.00	3.460	3.499	3.092	3.612	3.48	3.14
Σ Al to Mg	5.179	6.92	5.29	5.10	5.236	5.107	5.031	5.392	5.11	5.03
R ²⁺ in M4 site	0.179	1.92	0.29	0.10	0.236	0.107	0.031	0.392	0.11	0.03
Ca	1.754	0.10	0.39	1.76	1.659	1.633	1.654	1.479	1.57	1.66
Na	0.373	0.13	1.81	0.42	0.826	1.136	1.200	0.791	1.17	1.16
K	0.009	0.00	0.00	0.02	0.016	0.020	0.027	0.016	0.02	0.03

The structural formulae are calculated based on 23 oxygens and total Fe = Fe²⁺. R²⁺ = Mg + Fe + Mn. A, actinolite rimming hornblende. 1L, Cumingtonite lamellae; 2L, sodic amphibole lamellae; 3L, actinolite lamellae. 1H, Host hornblende of core; 2H, katophoritic outer core; 3H, Fe-enriched katophoritic rim. 1B, Host + lamellae of core; 2B, katophoritic outer core; 3B, Fe-enriched katophoritic rim.

The structural formulae of the clinoamphibole lamellae and host hornblendes, calculated from microprobe analyses on a basis of 23 oxygens and assuming total Fe to be ferrous, are listed in Table 1. The analyses of the lamellae were made at very low accelerating voltages with specimen currents of 0.01–0.02 μA (6–7 kV for Si, Al, Mg and Na, and 8–9 kV for Ca, Ti, Fe, Mn and K). Focusing spots in these conditions are just sufficient for resolution of lamellae 2.0-μm thick²⁴. Their compositions were calculated by comparing the intensity for each oxide component with those of the host hornblende after background and dead-time corrections. The resulting analyses must be of limited accuracy, yet they are sufficient to characterize the chemical properties of the lamellae.

The actinolite lamellae have a similar composition to that of the rimming actinolite. The A-site occupancy of 0.49 Na of the sodic amphibole lamellae is surprising as the A-sites of many metamorphic sodic amphiboles are almost empty¹⁶. This may be partly due to the limited accuracy of analysis, as described above, and partly a result of a substantially high Fe³⁺/(Fe²⁺ +

Fe³⁺) ratio in the amphibole²⁵. If the M4 site is occupied only by Ca and Na, and all Fe is presented as Fe³⁺, the A-site occupancy then decreases to 0.15 Na with a structural formulae of Na_{1.77}Ca_{0.38}Mg_{3.12}Fe_{0.00}Fe_{1.00}Al(VI)_{0.88}[Al(IV)_{0.44}Si_{7.56}]O₂₂(OH)₂. This rough estimation supposes that the sodic amphibole has a composition approximately intermediate between glaucophane and magnesioriebeckite.

The host hornblende compositions show no great variation in Ca and Al(IV) throughout the three distinct zones. However, there is a conspicuous difference in the assumed R²⁺ in their M4 site (where R²⁺ = Mg + Fe + Mn), with higher R²⁺ in the M4 site of the core than in the other two zones (Fig. 2). The bulk composition (host + lamellae) of the core was determined by analysis using a broad electron beam (30 μm in diameter), as the cumingtonite lamellae have an average spacing of ~5 μm. For the other two zones, the bulk compositions were estimated assuming the fractions of lamellae to be 5%, respectively, on the basis of modal analysis using back-scattered electron-scanning micrographs. The hornblende compositions thus obtained for the three zones are listed in Table 1. Chemical contrast among the three zones is then clearer than outlined above. The core has a composition of subcalcic hornblende having conspicuously high assumed R²⁺ in the M4 site. R²⁺ content in the M4 site, substituting for Ca + Na, is believed to depend on the forming temperature of hornblende, with hornblendes containing such a high cumingtonite component being restricted to amphibolite or higher grade rocks^{4,26}. The hornblende core composition might be a record of early high-temperature equilibration of this rock, predating the epidote amphibolite stage, but no evidence is available to confirm this.

The other two zones are katophoritic hornblende, and are expected to contain some sodic amphibole component. For the outer core, containing the sodic amphibole lamellae, if it is assumed that Ca + Na in M4 site is 2.0 then there is 0.44 Na in the M4 site with the maximum possible Fe³⁺/(Fe²⁺ + Fe³⁺) = 0.42 (ref. 25). The rim hornblende is chemically similar to the outer core, differing only in its relatively high total Fe content. From the fact that the exsolved amphibole in this rim is not sodic amphibole but actinolite, the rim hornblende may possibly have low Fe³⁺ content relative to the outer core. From the textures, the Fe enrichment of the rim is considered to have taken place when actinolite began to intergrow with the hornblende.

The following textural and chemical characters help to identify the development of the lamellae. (1) The extension of the cumingtonite lamellae towards the rim changing into the sodic amphibole and/or actinolite lamellae (Fig. 1). (2) The actinolite lamellae are continuous with the rimming actinolite, that is, mostly free from exsolution. (3) The actinolite lamellae have a similar composition to that of the rimming actinolite. Thus, the exsolution of actinolite must have closely postdated the time when the actinolite, omphacite and phengite formed at the

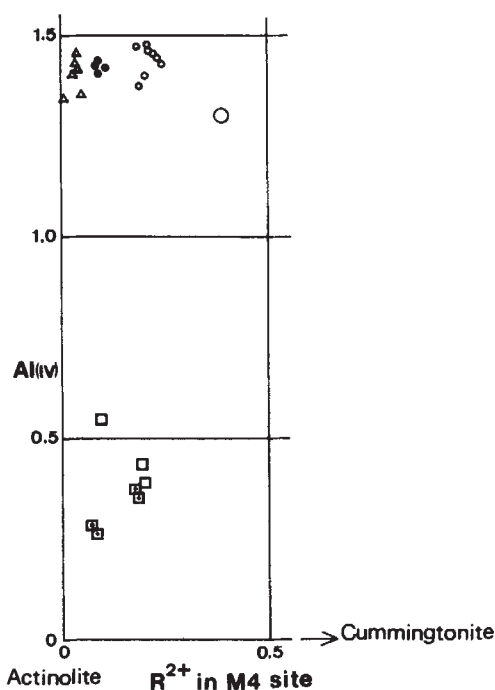


Fig. 2 Plot of R²⁺ in M4 site against Al(IV) for zoned hornblende rimmed by actinolite. Points analysed: ○, host of core; ●, host of katophoritic outer core; △, host of Fe-enriched katophoritic rim; ○, host + lamellae of core; □ and ■, actinolites, of which dotted squares represent analyses at points adjacent to the boundaries to hornblende.

glaucophane schist facies stage. It is thought that the exsolution also developed in the hornblende at the same time. The two inner zones of the hornblende could have become unstable rather earlier. The core hornblende with such a high cummingtonite component is probably unstable in the epidote amphibolite facies condition⁴. It is possible that the unmixing of the cummingtonite and the sodic amphibole preceded that of actinolite. Even if this is so the textures suggest that the present construction of the lamellae was completed at this retrograde stage. Cummingtonite and sodic amphibole still seem to be stable phases compared with the host, and the solvus of cummingtonite^{3,4} and sodic amphibole¹⁸ compared with calcic amphibole seems to extend into the glaucophane schist facies.

Our investigation is unfortunately restricted to the thick lamellae that are suitable for microprobe resolution. We can see a further set of finer microlamellae (Fig. 1), which are estimated to be cummingtonite from their back-scattered electron contrast. Further investigation of such fine and incipient exsolution texture is needed to explain fully the mechanism and history of the formation of the complicated lamellar texture and we cannot be sure that the exsolutions described here apply to stable equilibrium.

Many field studies have shown that calcic amphibole is barroisite or hornblende in the high-grade glaucophane schist facies^{7,8,11,16}, while at low-grade regimes it is actinolite with a miscibility gap to sodic amphibole¹⁸. The solvus in the hornblende-actinolite system has been thought to disappear in glaucophane schist parageneses^{7,11,16}. In contrast, the amphibole exsolution described here provides strong evidence indicating that this solvus extends into the high *P/T* metamorphic facies. A low K_D [= (total Fe/Mg)_{act} / (total Fe/Mg)_{horn}] value (0.64) of partitioning between the actinolite lamellae and the hornblende host suggests that the strongly Mg-preferring nature of actinolite in actinolite-hornblende pairs, previously known from many low-pressure metamorphic rocks¹⁵, is unchanged even in high-pressure conditions.

Hornblende and sodic amphibole associations have previously been described from high-grade glaucophane schist^{8,9} but were claimed to be non-equilibrium pairs and the conversion of hornblende to sodic amphibole is argued to be gradational¹⁶. Our investigation directly confirms a solvus between hornblende and sodic amphibole of glaucophane-magnesioriebeckite composition. This solvus may be extended to Fe²⁺-bearing sodic amphibole of riebeckite-crossite composition, and some of the previous two amphibole pairs, having chemically sharp two-phase boundaries^{8,9}, might be equilibrium pairs.

We thank S. Banno, T. Watanabe and A. F. Cooper for helpful discussions and critical reading of the manuscript.

Received 24 January; accepted 18 May 1983.

- Ghose, S. *Amphiboles and other Hydrous Pyroxenes—Mineralogy* (ed. Veblen, D. R.) 325–372 (Mineralogical Society of America, 1981).
- Robinson, P. et al. *Amphiboles: Petrology and Experimental Phase Relations* (eds Veblen, D. R. & Ribbe, P. H.) 1–227 (Mineralogical Society of America, 1982).
- Ross, M., Papike, J. J. & Weiblen, P. W. *Science* **159**, 1099–1104 (1968).
- Ross, M., Papike, J. J. & Wier Shaw, K. *Miner. Soc. Am. Spec. Pap.* **2**, 275–299 (1969).
- Tomita, K., Yamaguchi, Y. & Takita, R. *Mem. geol. Soc. Jap.* **11**, 95–106 (1974).
- Shido, F. & Miyashiro, A. *J. Fac. Sci. Univ. Tokyo* **12**, 85–102 (1959).
- Banno, S. *J. Fac. Sci. Univ. Tokyo* **15**, 203–319 (1964).
- Himmelberg, G. R. & Papike, J. J. *J. Petrol.* **10**, 102–114 (1969).
- Klein, C. *Am. Miner.* **54**, 212–237 (1969).
- Cooper, A. F. & Lovering, J. F. *Contr. Miner. Petrol.* **27**, 11–24 (1970).
- Ernst, W. G. *Geol. Soc. Am. Mem.* **135**, 73–94 (1972).
- Hietanen, A. *Am. Miner.* **59**, 22–40 (1974).
- Brady, J. B. *Am. Miner.* **59**, 529–535 (1974).
- Misch, P. & Rice, J. M. *J. Petrol.* **16**, 1–21 (1975).
- Tagiri, M. *Contr. Miner. Petrol.* **62**, 271–281 (1977).
- Ernst, W. G. *Miner. Mag.* **43**, 269–278 (1979).
- Spear, F. S. *Contr. Miner. Petrol.* **72**, 33–41 (1980).
- Maresch, W. V., Medenbach, O. & Rudolph, A. *Nature* **296**, 731–732 (1982).
- Oba, T. *Contr. Miner. Petrol.* **71**, 247–256 (1980).
- Shibakusa, H. *J. Geol. Jap.* **80**, 341–353 (in Japanese with English Abstr.) (1974).
- Shibakusa, H., Gouchi, N. & Imaizumi, M. *J. geol. Soc. Jap.* **83**, 301–303 (1977).
- Banno, S., Ishizuka, H. & Gouchi, N. *Abstr. Int. Geody. Conf.*, Tokyo, 14–15 (1978).
- Asahina, T. & Komatsu, M. *J. geol. Soc. Jap.* **85**, 317–330 (1979).
- Castaing, R. *Advances in Electronics and Electron Physics* Vol. 13, 317–386 (Academic, New York, 1960).
- Stout, J. H. *J. Petrol.* **13**, 99–145 (1972).
- Robinson, P. & Jaffe, H. W. *Miner. Soc. Am. Spec. Pap.* **2**, 251–274 (1969).

Interfacial sediment and assessment of organic input from a highly productive water column

Dedric J. Smith & Geoffrey Eglinton

Organic Geochemistry Unit, School of Chemistry,
University of Bristol, Bristol BS8 1TS, UK

Robert J. Morris

Institute of Oceanographic Sciences, Wormley, Godalming,
Surrey GU8 5UB, UK

Enormous quantities of organic matter are formed in the high-productivity areas of the world's oceans¹ (such as in upwelling areas), giving rise to underlying deposits of organic-rich sediments. Here, we have considered the problem of identifying the true organic input of these deposits by examining a sample of sediment from the sediment–water column interface of the Peru Upwelling. The fatty acid composition of this sediment can be rationalized in terms of a large phytoplankton input, plus a contribution from bacteria. The presence of significant amounts of labile polyunsaturated acids implies that the phytoplankton lipids have not suffered substantial chemical or biological breakdown. This suggests that in such sedimentary regimes the first few millimetres of material lying at the sediment–water column interface are representative of the actual material reaching the sediments, and offer a useful guide to the true organic input.

A full understanding of sedimentation processes requires data on the original sources of the sedimentary material (mainly phytoplankton in areas of very high productivity), the sedimentary particulates and the sediments themselves in various stages of diagenesis. Numerous analyses have been made on phytoplankton species^{2–12}, sediment trap particulates^{13–17} and bulk sediment^{18–23}. A valuable contribution to such work may be made by analysis of the material found at the interface of the sediment with the water column. This unconsolidated material has completed its passage through the water column, but has not yet become incorporated into the bulk sediment. By comparison with the likely original sources of sedimentary lipids, its chemical composition should give an indication of the alterations which occur during sedimentation, while also serving as a baseline for assessing post-depositional diagenetic changes in older subsurface sedimentary samples. In unconsolidated sediments, processes such as resuspension of material by current action may make it difficult to define just where the sediment–water column interface occurs. Similarly, part of the surficial sediment may be lost during sampling. Nevertheless, we feel that careful use of coring devices specifically designed to cause minimum disturbance of the sediment surface²⁴, enables collection of a realistic sample from this important region.

The sample discussed here was taken by carefully removing the loose material from the surface of the sediment (1–2 mm layer) in a box core²⁵. The core was taken in June 1980 during Cruise 110 of RRS *Discovery*, from a position 12°01.8' S, 77°29.3' W, in a water depth of 145 m, on the Peruvian continental shelf. The interface consisted of a dark green, fluid diatomaceous ooze which was highly anoxic.

On the basis of sedimentation rates calculated for this area from sediment trap data¹⁶, and sediment ¹⁴C data (D. Harkness, personal communication), the sample is thought to be <1 yr old. Extraction and analytical procedures are described elsewhere²⁵. The composition of the fatty acid fraction, obtained after saponification of the total solvent extract, is shown in Fig. 1. The total organic carbon content of the sediment was 16.9% (dry weight) and the total extractable lipid amounted to 4.5% (dry weight).

The most notable feature of the fatty acid distribution is the presence of an appreciable polyunsaturated component. Polyunsaturated fatty acids are common among many species of organism, such as phytoplankton and zooplankton, which are believed to be major contributors of organic matter to sediments^{2-12,16,17}; 20:5 and 22:6 have been found in sediment traps at a depth of 52 m in the Peru Upwelling region¹⁷, indicating that transport of polyunsaturates on particulates does occur through the water column. However, these highly labile compounds have rarely been reported to occur in sediments²¹⁻²³, presumably due to their rapid degradation in the water column and/or surface sediment. Their presence in this sample suggests that the original fatty acids and, by implication, other natural product lipids, have suffered little degradation since their original biosynthesis, and therefore this interfacial material may provide an accurate reflection of the sources of sedimentary organic matter.

Mineralogical and microscopic data on this sediment suggest an input dominated by diatoms, with a lesser contribution from silicoflagellates²⁵. Many diatom cells visible in the sample were intact and still contained pigments in their chloroplasts²⁵. Unfortunately, a sample of the plankton of the area was not

available for comparison. However, a considerable number of marine phytoplankton species have been analysed for fatty acids, and Tables 1 and 2 present data for 43 such species, plus two mixed cultures. Detailed comparisons are not really appropriate for several reasons: the fatty acid composition of organisms may be dependent on culture conditions^{10,26}, while limitations of early packed column gas chromatography may have produced inaccuracies in some quantitative data due to non-resolution of components. Thus, some early analyses did not report isomer compositions for some acids where multiple isomers might be expected^{4,9}. Tables 1 and 2 are designed to provide a broad comparison of the major phytoplankton groups, and certain general chemotaxonomic distinctions can be seen, particularly among the C₁₆, C₁₈ and C₂₀ acids.

The Peru interfacial sediment contains high levels of 16:0 and 16:1 acids and low levels of C₁₈ acids, and in these respects matches the diatoms more closely than the other groups. The major sedimentary polyunsaturated acid is 20:5, again matching most diatom species, although the levels in the sediment are lower than in phytoplankton. This is true of all the polyunsaturated fatty acids and can be attributed to faster degradation of the polyunsaturated components compared with the saturates and monounsaturates. The Walvis Bay data (to be published in detail elsewhere) are included for comparative purposes; this sediment is an organic-rich diatomaceous ooze with a very similar mineralogy to the Peru sample²⁷, but older; ~70 yr. Polyunsaturated acids were not detected in this sediment, and monounsaturates are at much lower levels compared with the saturates than in the Peru sediment, implying much more extensive degradation of the acids.

Differences in the proportions of some acids in the sediment, as compared with those of the phytoplankton, are noteworthy. The ratio of 20:4/20:5 is higher in the former, possibly due to preferential removal of 20:5. Other sources of acids may exist, although the sediment is certainly dominated by phytoplankton remains²⁸. Zooplankton, especially calanoid copepods, may be significant contributors of sedimentary lipids in this area^{16,17}, but frequently show a fatty acid composition similar to that of their phytoplankton diet^{29,30}. In any case, the very low levels of wax esters and fatty alcohols found in this sediment sample²⁵ would seem to preclude a major calanoid copepod contribution¹⁷.

The 22:6 acid, however, is comparatively more abundant in the sediment than it is in the diatoms; the phytoplankton containing the highest concentrations of 22:6 are certain dinoflagellate species. Although no dinoflagellate cysts were visible in the sample, there is evidence from the sterol distribution of a 'dinoflagellate-like' input and the occurrence of 22:6 may be related to this²⁸.

Although 18:1 Δ^9 is of low abundance, in accord with the diatom data, 18:1 Δ^{11} is relatively high. Unlike phytoplankton, bacteria tend to produce 18:1 Δ^{11} in much greater amounts than the Δ^9 isomer: this observation has been used to characterize bacterial contributions to sediments^{22,23}. In addition, branched, odd-chain acids, such as the iso- and anteiso-C₁₃, C₁₅ and C₁₇ compounds present in this sample, are commonly taken as bacterial markers^{19,22,23}. Hence, we consider that most of the 18:1 Δ^{11} , together with the odd-carbon branched chain acids and some of the other saturates and monounsaturates are derived from bacteria. These microorganisms, by utilizing part of the sedimentary organic material, may introduce changes in the general lipid composition. Such reworking can produce a more saturated suite of fatty acids^{31,32}, as may have occurred with the Walvis Bay sediment¹⁹. Since the polyunsaturated fatty acids in this sample are characteristic of many plankton species²⁻¹², but are rare among marine bacteria^{22,33,34}, the original biological input to the sediment does not appear to have been significantly altered by metabolic microbiological reworking. The most notable loss, not unexpectedly, has been in the labile, highly unsaturated components, such as the 20:5 fatty acid.

The fatty acids of this Peruvian interfacial sediment are thus

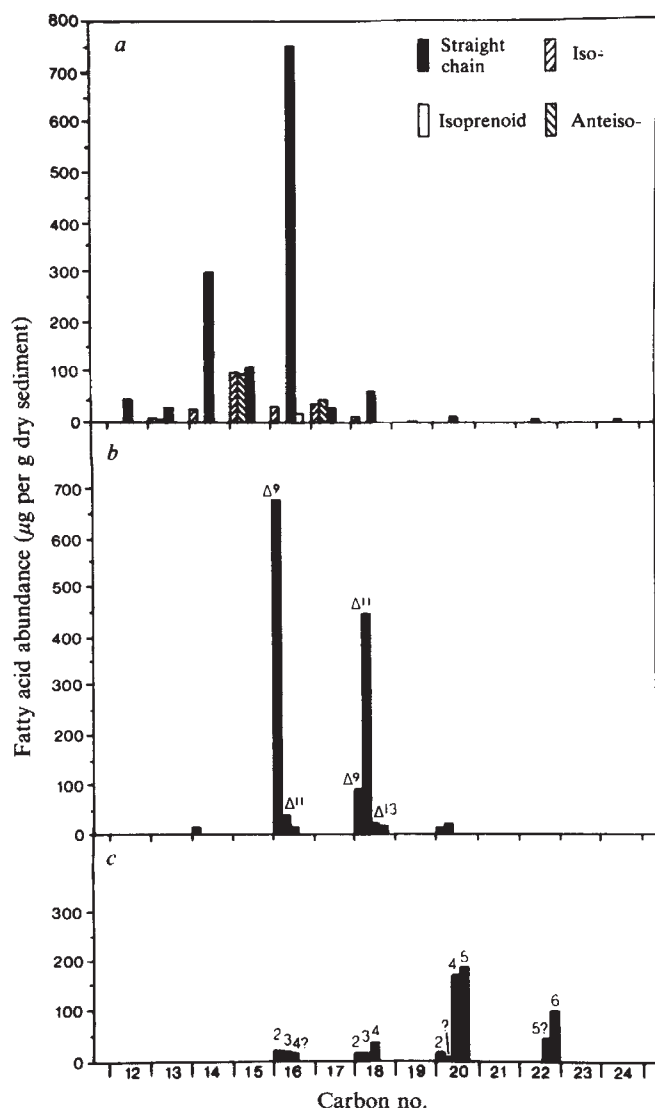


Fig. 1 Fatty acid distribution of an interfacial sediment sample from the Peru Upwelling. Acids obtained after alkaline hydrolysis of the extracted lipids. *a*, Saturated fatty acids. *b*, Monounsaturated fatty acids. The double bond position for the major isomers is given above the histogram bar. *c*, Polyunsaturated fatty acids. The number of double bonds is given above the histogram bar.

Table 1 Major fatty acids of marine surface sediments and of phytoplankton—relative abundances (%)

Sediment sample/ phytoplankton species†	Major fatty acid abundances*											Refs	
	12:0	14:0	16:0	16:1	16:μ‡	18:0	18:1Δ ⁹	18:1Δ ¹¹	18:μ‡	20:4	20:5		22:6
Sediment													
Peru interface	<5	5–10	>20	>20	<5	<5	<5	10–15	<5	<5	5–10	<5	Fig. 1
Walvis Bay 0–2 cm	<5	10–15	>20	5–10	—	5–10	<5	<5	—	—	—	—	19,§
Bacillariophyceae													
<i>Amphora</i> sp.	—	5–10	15–20	>20	?	—	?	?	<5	<5	10–15	—	9
<i>Asterionella japonica</i>	<5	5–10	10–15	>20	10–15	<5	<5	<5	<5	5–10	15–20	—	4
<i>Biddulphia</i> sp.	—	15–20	10–15	>20	10–15	<5	<5	<5	<5	<5	10–15	—	4, 9, 11
<i>Chaetoceros septentrionalis</i>	<5	5–10	10–15	15–20	10–15	<5	<5	<5	<5	<5	>20	—	4
<i>Coscinodiscus eccentricus</i>	<5	10–15	15–20	15–20	15–20	<5	<5	—	5–10	<5	5–10	<5	10
<i>Cyclotella cryptica</i>	<5	5–10	>20	>20	5–10	<5	<5	<5	<5	<5	10–15	<5	2, 6
<i>Cylindrotheca</i> sp.	—	5–10	15–20	>20	5–10	<5	<5	<5	<5	5–10	>20	—	5, 6
<i>Cylindropiscis</i> sp.	—	15–20	15–20	>20	5–10	<5	<5	—	<5	<5	10–15	—	5
<i>Ditylum brightwellii</i>	<5	5–10	10–15	>20	10–15	<5	5–10	<5	<5	<5	10–15	—	4
<i>Fragilaria</i> sp.	—	>20	15–20	>20	?	—	?	?	5–10	—	10–15	<5	9
<i>Lauderia borealis</i>	—	5–10	10–15	>20	10–15	—	<5	<5	<5	<5	>20	—	4
<i>Navicula</i> sp.	—	<5	15–20	>20	15–20	<5	5–10	—	<5	<5	15–20	—	5, 6, 9
<i>Nitzschia</i> sp.	—	5–10	>20	>20	5–10	<5	<5	<5	<5	<5	5–10	—	5, 6, 9
<i>Phaeodactylum tricornutum</i>	<5	5–10	15–20	>20	10–15	<5	<5	<5	<5	<5	15–20	—	2, 4, 6, 7
<i>Skeletonema costatum</i>	<5	15–20	10–15	>20	15–20	<5	<5	—	<5	<5	15–20	<5	2, 4
<i>Thalassiosira</i> sp.	<5	10–15	15–20	>20	5–10	<5	<5	<5	<5	5–10	10–15	<5	2, 9
Mixed diatom culture	<5	<5	15–20	>20	5–10	<5	<5	<5	<5	5–10	15–20	—	34
Haptophyceae													
<i>Chrysalolithus hyalinus</i>	<5	10–15	>20	<5	<5	<5	<5	<5	>20	—	5–10	5–10	12
<i>Dicrateria inornata</i>	<5	<5	15–20	5–10	10–15	—	15–20	<5	>20	<5	5–10	—	4
<i>Emiliania huxleyi</i>	<5	>20	5–10	<5	<5	<5	10–15	<5	>20	—	<5	10–15	4, 12
<i>Hymenomonas</i> sp.	<5	<5	15–20	10–15	10–15	<5	5–10	<5	>20	<5	10–15	<5	4, 12
<i>Isochrysis galbana</i>	<5	>20	10–15	<5	<5	<5	10–15	<5	>20	—	<5	10–15	12
<i>Pavlova lutheri</i>	<5	10–15	10–15	>20	15–20	<5	<5	<5	<5	<5	15–20	10–15	2, 4
<i>Prymnesium parvum</i>	—	<5	10–15	5–10	<5	<5	>20	<5	>20	<5	<5	—	4, 5
Chlorophyceae													
<i>Chlamydomonas</i> sp.	<5	<5	15–20	<5	<5	<5	15–20	<5	>20	—	—	—	4, 5
<i>Chlorella prototheca</i>	—	<5	>20	<5	15–20	<5	5–10	<5	>20	—	—	—	5
<i>Dunaliella</i> sp.	<5	<5	10–15	5–10	15–20	<5	<5	<5	>20	<5	5–10	—	2, 4
Mixed green algae culture	<5	<5	>20	5–10	10–15	<5	10–15	<5	>20	<5	<5	—	34
Dinophyceae													
<i>Amphidinium carterii</i>	<5	<5	>20	<5	<5	5–10	5–10	<5	10–15	—	5–10	>20	2
<i>Gonyaulax polyedra</i>	—	<5	>20	<5	<5	<5	<5	<5	15–20	—	10–15	>20	4
<i>Peridinium trochoideum</i>	<5	15–20	>20	<5	<5	—	5–10	<5	10–15	<5	10–15	—	4
<i>Prorocentrum micans</i>	<5	10–15	5–10	>20	>20	—	5–10	<5	5–10	—	10–15	—	4

Abundances are not precise values, and are expressed as ranges only to give a qualitative 'fingerprint' suitable for consideration of overall similarities or differences.

* All data are for the fatty acids obtained after alkaline hydrolysis of the lipids expressed as % of total fatty acids. —, Not reported. ?, Abundance uncertain due to co-elution of 16:3 and 18:1.

† Phytoplankton classified according to Parke and Dixon³⁵ and Hendly³⁶. Species belonging to a common genus have been grouped under the generic name as follows: *Amphora*: *A. exigua* and unspecified species. *Biddulphia*: *B. aurita*, *B. sinensis*. *Cylindrotheca*: *C. fusiformis*, *C. gracilis*. *Cylindropiscis*: unspecified species. *Fragilaria*: unspecified species. *Navicula*: *N. abscondita*, *N. biskantieri*, *N. pelliculosa*. *Nitzschia*: *N. frustulum*, *N. longissima*, *N. ovalis*, *N. thermalis*. *Thalassiosira*: *T. fluviatilis*, *T. pseudonana*. *Hymenomonas*: *H. carterae*, *H. elongata*. *Chlamydomonas*: *C. palla* and unspecified species. *Dunaliella*: *D. primolecta*, *D. tertiolecta*. The data from multiple analyses have been averaged.

‡ Polyunsaturated C₁₆ or C₁₈ acids.

§ Unpublished data.

// Isomer composition not reported.

Table 2 Abundances of major fatty acids reported for marine surface sediments and phytoplankton

Sediment/phytoplankton group	Major fatty acid abundances*											
	12:0	14:0	16:0	16:1	16:μ [†]	18:0	18:1Δ ⁹	18:1Δ ¹¹	18:μ [†]	20:4	20:5	22:6
Sediment												
Peru interface	1.3	8.9	22.2	20.4	1.7	1.8	2.7	12.9	2.0	4.9	5.5	2.9
Walvis Bay 0–2 cm	0.3	12.2	68.7	6.0	—	9.6	2.0	2.8	—	—	—	—
Phytoplankton†												
Bacillariophyceae	0–1	2–32	2–38	13–56	5–20	0–3	1–7	0–2	0–8	0–14	4–44	0–4
Haptophyceae	0–1	1–35	5–26	1–24	1–20	0–4	4–25	0–2	3–45	0–4	2–22	0–14
Chlorophyceae	0–1	1–4	9–21	2–10	3–21	0–4	1–12	1–3	25–48	0–4	0–10	—
Dinophyceae	0–1	2–17	10–36	1–23	1–25	0–6	1–7	0–1	8–18	0–1	7–14	0–25

* Abundances expressed as % of total fatty acids. For phytoplankton, the range of the abundances reported for the species shown in Table 1 is given.

† Polyunsaturated C₁₆ or C₁₈ acids.

‡ Phytoplankton species given in Table 1.

believed to come principally from the diatom blooms known to occur in the surface waters, and have suffered only partial degradation of the polyunsaturated components. Sedimentation in this area is thought to occur mainly as a consequence of such productivity 'events'; the rapid sinking of the blooms results in large pulses of organic matter being quickly transported to the underlying sediments in a largely undegraded state. The fatty acids should be a valid representation of the organic matter in the initial stages of sediment formation. Since these readily-metabolizable compounds are not extensively degraded, the

same is likely to be true of other lipid classes, and these are reported elsewhere^{25,28}. We believe that such data, in conjunction with similar information obtained from samples representative of organic matter, both before deposition and after burial in the sediment column, will provide a basis for understanding the initial diagenesis of natural product organic compounds.

We acknowledge the NERC for providing funding for gas chromatography-mass spectrometry facilities (grants GR3/2951 and GR3/3758), and a studentship (GT4/79/ALS/5) for D.J.S.

Received 18 April; accepted 12 May 1983.

1. Demaison, G. J. & Moore, G. T. *Organ. Geochem.* **2**, 9–31 (1980).
2. Ackman, R. G., Tocher, C. S. & McLachlan, J. J. *Fish. Res. Bd Can.* **25**, 1603–1620 (1968).
3. Chu, F.-L. E. & Dupuy, J. L. *Lipids* **15**, 356–364 (1980).
4. Chuecas, L. & Riley, J. P. *J. mar. biol. Ass. U.K.* **49**, 97–116 (1969).
5. De Mort, C. L., Lowry, R., Tinsley, I. & Phinney, H. K. *J. Phycol.* **8**, 211–216 (1972).
6. Kates, M. & Volcani, B. E. *Biochim. biophys. Acta* **116**, 264–278 (1966).
7. Moreno, V. J., De Moreno, J. E. A. & Brenner, R. R. *Lipids* **14**, 15–19 (1979).
8. Morris, R. J. & Culkin, F. *Oceanogr. mar. Biol. A. Rev.* **14**, 391–433 (1976).
9. Orcutt, D. M. & Patterson, G. W. *Comp. Biochem. Physiol.* **50B**, 579–583 (1975).
10. Pugh, P. R. *Mar. Biol.* **11**, 118–124 (1971).
11. Volkman, J. K., Eglinton, G. & Corner, E. D. S. *Phytochemistry* **19**, 1809–1813 (1980).
12. Volkman, J. K., Smith, D. J., Eglinton, G., Forsberg, T. E. V. & Corner, E. D. S. *J. mar. biol. Ass. U.K.* **61**, 509–527 (1981).
13. Wiebe, P. H., Boyd, S. H. & Winget, C. J. *mar. Res.* **34**, 341–354 (1976).
14. Prahl, F. G., Bennett, J. T. & Carpenter, R. *Geochim. cosmochim. Acta* **44**, 1967–1976 (1980).
15. Wakeham, S. G. *et al.* *Nature* **286**, 798–800 (1980).
16. Gagosian, R. B., Volkman, J. K. & Nigrelli, G. E. in *Advances in Organic Geochemistry 1981* (eds Bjorøy, M. *et al.*) (Wiley, London, in the press).
17. Wakeham, S. G., Farrington, J. W. & Volkman, J. K. in *Advances in Organic Geochemistry 1981* (eds Bjorøy, M. *et al.*) (Wiley, London, in the press).
18. Morris, R. J. & Calvert, S. E. in *A Voyage of Discovery: George Deacon 70th Anniversary Volume* (ed. Angel, M.) 647–680 (Pergamon, Oxford, 1977).
19. Boon, J. J., De Leeuw, J. W. & Schenck, P. A. *Geochim. cosmochim. Acta* **39**, 1559–1565 (1975).
20. Farrington, J. W. & Quinn, J. G. *Geochim. cosmochim. Acta* **37**, 259–268 (1973).
21. Kawamura, K., Ishiwatari, R. & Yamazaki, M. *Chem. Geol.* **28**, 31–39 (1980).
22. Perry, G. J., Volkman, J. K., Johns, R. B. & Bavor, H. J. *Geochim. cosmochim. Acta* **43**, 1715–1725 (1979).
23. Volkman, J. K. & Johns, R. B. *Nature* **267**, 693–694 (1977).
24. Peters, R. D., Timmins, N. T., Calvert, S. E. & Morris, R. J. *I.O.S. Rep. No. 106*, ((1980).
25. Smith, D. J., Eglinton, G. & Morris, R. J. *Geochim. cosmochim. Acta* (in the press).
26. Wright, D. C., Berg, L. R. & Patterson, G. W. *Phytochemistry* **19**, 783–785 (1980).
27. Smith, D. J., Eglinton, G., Morris, R. J. & Poutanen, E. L. *Oceanol. Acta* **5**, 365–378 (1982).
28. Smith, D. J., Eglinton, G., Morris, R. J. & Poutanen, E. L. *Oceanol. Acta* **6**, 211–219 (1983).
29. Moreno, V. J., De Moreno, J. E. A. & Brenner, R. R. *Oceanol. Acta* **2**, 373–378 (1979).
30. Volkman, J. K., Corner, E. D. S. & Eglinton, G. in *Colloques int. Cent. natn. Rech. scient.* No. 293, 185–197 (1979).
31. Rhead, M. M., Eglinton, G., Draffan, G. H. & England, P. J. *Nature* **232**, 327–330 (1971).
32. Gaskell, S. J., Rhead, M. M., Brooks, P. W. & Eglinton, G. *Chem. Geol.* **17**, 319–324 (1976).
33. Oliver, J. D. & Colwell, R. R. *Int. J. Syst. Bact.* **23**, 442–458 (1973).
34. Volkman, J. K., Johns, R. B., Gillan, F. T., Perry, G. J. & Bavor, H. J. *Geochim. cosmochim. Acta* **44**, 1133–1143 (1980).
35. Parke, M. & Dixon, P. S. *J. mar. biol. Ass. U.K.* **56**, 527–594 (1976).
36. Hendy, N. I. *J. mar. biol. Ass. U.K.* **54**, 277–300 (1974).

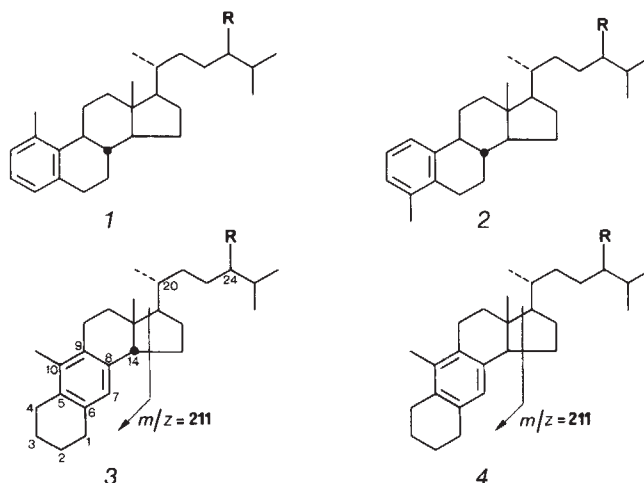


Fig. 1 Major series of monoaromatic steroid hydrocarbons occurring in immature Cretaceous black shales from the DSDP. R=H; CH₃; C₂H₅.

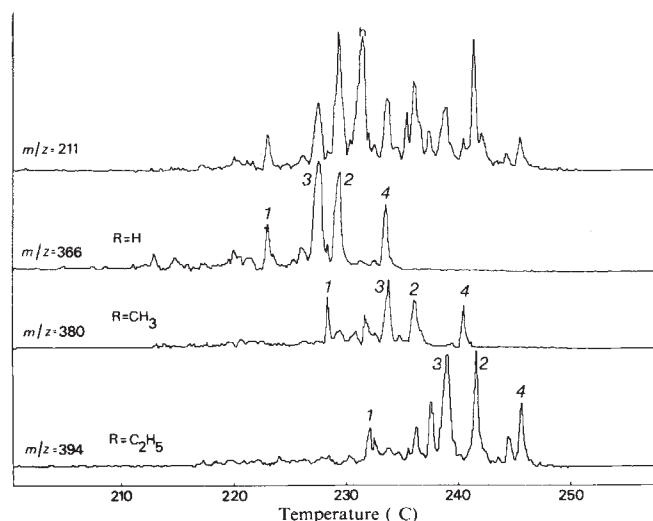


Fig. 2 Mass fragmentograms of ions 211, 366, 380 and 394 showing the distribution of ring A monoaromatic steroid hydrocarbons in a Cretaceous black shale from the DSDP (Leg 40, Site 364). Numbers refer to structures in Fig. 1. *h*, Monoaromatic C₂₇ hopanoid. Conditions: SP 2250, 25 m × 0.25 mm, 0.08 μm, 140–250 °C, 2 °C min⁻¹. Computerized GC–MS, LKB 9000 S.

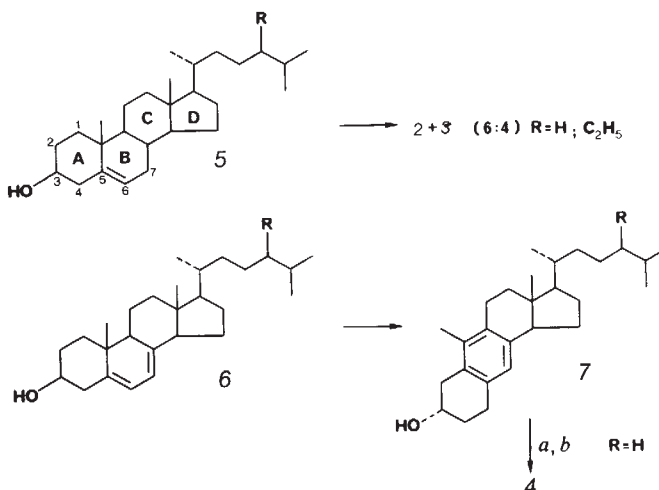


Fig. 3 Scheme of synthesis of anthrasteroid hydrocarbons 3 and 4, after refs 12 and 13. *a*, H₂SO₄, acetone; *b*, N₂H₄; KOH, (CH₂OH)₂.

C₂₇–C₂₉ Monoaromatic anthrasteroid hydrocarbons in Cretaceous black shales

G. Hussler & P. Albrecht

Laboratoire de Chimie Organique des Substances Naturelles, associé au CNRS, Département de Chimie, Université Louis Pasteur, 1, rue Blaise Pascal, Strasbourg, France

The structural elucidation of aromatic steroid hydrocarbons occurring in sediments and petroleum has been the subject of increasing interest during recent years^{1–7}. Aromatization of the precursor sterols of marine origin has been shown to start at early stages of diagenesis in immature sediments^{8,9}. In elucidating these first aromatization reactions we have recently identified two major series of ring A monoaromatic steroids, 1 and 2 (Fig. 1), in Cretaceous black shales from the southern and eastern North Atlantic, collected during the Deep Sea Drilling Project (DSDP)⁸. We report here the identification of two novel series of geochemical markers, the ring B monoaromatic anthrasteroid hydrocarbons 3 and 4, as the two other major series occurring in the same immature DSDP samples (Fig. 1).

The shale samples and the separation scheme of the mono- + diaromatic hydrocarbons have been described previously⁸. Ring A or B aromatic steroid hydrocarbons display a major fragmentation at *m/z* = 211. Figure 2 shows a mass fragmentogram of ion 211 corresponding to the distribution of this class of compounds, as well as the fragmentograms of the molecular ions *m/z* = 366, 380 and 394 corresponding respectively to the C₂₇, C₂₈ and C₂₉ compounds.

The skeleton rearrangement leading to the anthrasteroid series 3 and 4 has already been well studied. The two isomers differing only in their stereochemistry at position 14 have been

obtained from synthesis and depend on the precursor and the reaction conditions^{10,11}.

Compounds 3 ($R=H$ and $R=C_2H_5$), 14 β (H)-1(10 $\rightarrow$ 6)-abeo-cholesta-5,7,9(10)-triene and 24(R)ethyl-14 β (H)-1(10 $\rightarrow$ 6)-abeo-cholesta-5,7,9(10)-triene were prepared respectively from cholest-5-en-3 β -ol (cholesterol; 5, $R=H$) and 24(R)ethyl-cholest-5-en-3 β -ol (sitosterol; 5, $R=C_2H_5$) as starting materials, as previously described for lower homologues¹² (Fig. 3). Compounds 3 were separated from the 4-methyl-19-nor-cholesta-1,3,5(10)-trienes, 2, obtained in the same procedure, by reversed phase HPLC (RP 18) using methanol as the mobile phase.

Compound 4 ($R=H$) was prepared using a method which does not affect the stereochemistry at position 14 during the rearrangement of the steroid skeleton¹³. In this procedure 14 α (H)-1(10 $\rightarrow$ 6)-abeo-cholesta-5,7,9(10)-triene was obtained from cholesta-5,7-dien-3 β -ol, 6, as a starting material (Fig. 3).

Spectrometric data observed for compounds 3 and 4 agree with those given in the literature^{10,14} and were confirmed by proton magnetic resonance and mass spectrometry (MS). The chemical shift of the C-18 methyl group can be used to distinguish the two isomers¹⁴. Indeed when the C/D ring junction is *trans*, as in compound 4, the C-18 methyl group is located in the shielding zone of the aromatic ring ($\delta = 0.59$ p.p.m.). When the C/D ring junction is *cis*, as in compound 3, this methyl group is withdrawn from the influence of the aromatic ring and appears at lower field ($\delta = 0.98$ p.p.m.).

3 $R=H$ 14 β (H)-1(10 $\rightarrow$ 6)-abeo-cholesta-5,7,9(10)-triene
NMR (200 MHz; $CDCl_3$; δ p.p.m.; J (Hz): 0.87 (6H,d, J = 6.5); 0.97 (3H,d, J = 5.5); 0.98 (3H,s); 2.12 (3H,s); 6.78 (1H,s). MS (70 eV): m/z = 366(M^+ , 100%); 351(5%); 253(9%); 226(34%); 211(80%); 199(27%); 197(19%); 170(4%); 169(14%); 159(15%); 158 (5%); 157(9%); 143(7%).

3 $R=C_2H_5$ 24(R)-ethyl-14 β (H)-1(10 $\rightarrow$ 6)-abeo-cholesta-5,7,9(10)-triene
NMR (250 MHz; $CDCl_3$; δ p.p.m.; J (Hz): 0.82 (3H,d, J = 5.7); 0.85 (3H,d, J = 5.7); 0.86 (3H,t, J = 5.7); 0.97 (3H,d, J = 5.7); 0.97 (3H,s); 2.07 (3H,s); 6.63 (1H,s). MS (70 eV): m/z = 394(M^+ , 100%); 379(4%); 253(9%); 226(33%); 211(71%); 199(28%); 197(16%); 170(3%); 169(11%); 159(13%); 158(5%); 157(7%); 143(6%).

4 $R=H$ 14 α (H)-1(10 $\rightarrow$ 6)-abeo-cholesta-5,7,9(10)-triene
NMR (200 MHz; $CDCl_3$; δ p.p.m.; J (Hz): 0.59 (3H,s); 0.88 (6H,d, J = 6.5); 0.99 (3H,d, J = 6.2); 2.07 (3H,s); 6.63 (1H,s). MS (70 eV): m/z = 366(M^+ , 100%); 351(12%); 253(17%); 226(33%); 211(81%); 199(56%); 197(16%); 170(5%); 169(12%); 159(23%); 158(4%); 157(9%); 143(8%).

The mass spectra of the *cis* and *trans* series are similar except for a significant difference in the intensity of the ions at m/z = 199 and 253. They are, however, easily distinguished from the ring A monoaromatic steroids 1 and 2. Indeed in compounds 3 and 4 the base peak is the molecular ion, whereas it is the fragment at m/z = 211 in 1 and 2.

As reported previously⁸ for compounds 1 and 2, compounds 3 ($R=H$; C_2H_5) and 4 ($R=H$) were identified in the shale samples by comparison with the authentic standards on the following basis: identical mass spectra and co-elution on three glass capillary columns (OV 73; SP 2250, Supelco; Pluronic F 68; respectively 30 m, 20 m and 50 m $\times$ 0.25 mm). Compounds 3 ($R=CH_3$) and 4 ($R=CH_3$; C_2H_5) were tentatively identified on the basis of their mass spectra and retention times.

These results demonstrate the occurrence of another aromatization pathway undergone by sterols during early stages of diagenesis in the subsurface, that is aromatization starting in ring B, after a rearrangement of the steroid skeleton, rather than in rings A or C^{8,2,15}. As the monoaromatic anthrasteroids are formed in the laboratory from the same precursor sterols in the same acidic conditions as ring A monoaromatic steroids¹², it is likely that this is also the case in the subsurface where the acid catalysis could be provided by clay minerals. This catalytic effect was indeed shown to operate in laboratory simulation

experiments in which cholesterol, heated in mild conditions in the presence of montmorillonite, led to a mixture containing both ring A and anthra monoaromatic steroid hydrocarbons (Sieskind and P. A., unpublished data). Like the ring A monoaromatic steroids, the anthrasteroids have not yet been observed in significant amounts in mature sediments and crude oils. Both classes of compounds seem to be transient species, limited to early stages of diagenesis, which are transformed during further maturation into triaromatic steroids⁴ or other phenanthrene or anthracene-type hydrocarbons.

Preliminary results on several DSDP samples from Legs 40 and 41⁸ seem to indicate that the relative ratio of compounds 3 and 4 could be a sensitive maturation indicator of slightly diagenetized sediments, in a zone where most other criteria used for more mature samples are still inoperative¹⁶. Indeed epimerization of the benzylic position at C-14 leading to the more stable *cis* C/D ring junction¹⁴ of compound 3 obviously takes place readily in the subsurface, as shown by the increasing proportions of the latter in the deeper, slightly more mature samples.

We thank the Institut Français du Pétrole for financial support and supply of samples, also Professeur G. Ourisson for his interest and helpful discussions.

Received 5 April; accepted 9 May 1983.

1. Spyckelle, C. thesis, Univ. Louis Pasteur (1975).
2. Schaeffe, J. thesis, Univ. Louis Pasteur (1977).
3. Schaeffe, J., Ludwig, B., Albrecht, P. & Ourisson, G. *Tetrahedron Lett.* 4163–4166 (1978).
4. Ludwig, B., Hussler, G., Wehrung, P. & Albrecht, P. *Tetrahedron Lett.* 22, 3313–3316 (1981).
5. Mackenzie, A. S., Hoffman, C. F. & Maxwell, J. R. *Geochim. cosmochim. Acta* 45, 1345–1355 (1981).
6. Seifert, W. K., Carlson, R. M. & Moldowan, J. M. in *Advances in Organic Geochemistry 1981* (eds Björroy, M. et al.) 710–724 (Wiley, London, 1983).
7. Mackenzie, A. S., Brassell, S. C., Eglinton, G. & Maxwell, J. R. *Science* 217, 491–504 (1982).
8. Hussler, G., Chappe, B., Wehrung, P. & Albrecht, P. *Nature* 294, 556–558 (1981).
9. Rullkötter, J., von der Dick, H. & Welte, D. H. *Init. Rep. DSDP Leg 64*, 837–853 (1982).
10. Tsuda, K., Hayatsu, R., Steele, A. J., Tanaka, O. & Mosettig, E. *J. Am. chem. Soc.* 85, 1126–1131 (1963).
11. Tanaka, O. & Mosettig, E. *J. Am. chem. Soc.* 85, 1131–1133 (1963).
12. Ogilvie, A. G. & Hanson, J. R. *JCS Perkin I*, 1981–1983 (1972).
13. Bosworth, N. et al. *JCS Perkin I*, 805–809 (1977).
14. Steele, A. S., Cohen, L. A. & Mosettig, E. *J. Am. chem. Soc.* 85, 1134–1138 (1963).
15. Mackenzie, A. S., Lamb, N. A. & Maxwell, J. R. *Nature* 295, 223–226 (1982).
16. Seifert, W. K. & Moldowan, J. M. *Geochim. cosmochim. Acta* 42, 77–92 (1978).

⁴⁰Ar/³⁹Ar age limit for an Acheulian site in Israel

G. Feraud*§, D. York*, C. M. Hall*, N. Goren† & H. P. Schwarcz‡§

* Department of Physics, University of Toronto, Toronto, Canada M5S 1A7

† Institute of Archaeology, Hebrew University of Jerusalem, Jerusalem, Israel

‡ Department of Geology, McMaster University, Hamilton, Canada L8S 4L8

The lake of Berekhat Ram fills a caldera which cuts the Pleistocene basalt flows that cover the Golan Heights¹ in Israel (Fig. 1). In the north-west wall of the crater a sequence of two basalt flows is exposed separated by a reddish-brown palaeosol complex ~2 m thick. Scattered through this palaeosol, and also found in prominent concentration in one layer near the base, are lithic artefacts of an Upper Acheulian industry. We have now obtained a ⁴⁰Ar/³⁹Ar date of 233 $\pm$ 3 kyr for the overlying flow, while the underlying flow does not give a well-defined plateau during step heating, but gives a range of apparent ages from 290 to 780 kyr, with an integrated age of 470 kyr.

Only the upper 2 m of the underlying basalt is exposed. This is a vesicular, holocrystalline aggregate of fine plagioclase laths in which clusters of augite crystals and skeletal olivine are set.

§ Present addresses: Laboratoire de Géologie et Géochemie, Faculté des Sciences, Parc Valrose, 06034 Nice Cedex, France (G.F.); Department of Geology, Hebrew University of Jerusalem, Jerusalem, Israel (H.P.S.).

Table 1 Conventional K-Ar analyses

Sample	K (%)	Mass (g)	Atmospheric contamination (%)	^{40}Ar radiogenic ($\times 10^{-13} \text{ mol g}^{-1}$)	Age (kyr)
BRG/81/23	1.743	3.3247	76.3	7.606	252 ± 14
BRG/81/26	0.844	3.0490	89.5	7.469	510 ± 32

Decay and isotopic constants as suggested by Steiger and Jäger¹².

Table 2 $^{40}\text{Ar}/^{39}\text{Ar}$ analytical data

Temperature (°C)	Atmospheric contamination (%)	$^{37}\text{Ar}_{\text{Ca}}/^{39}\text{Ar}_{\text{K}}$	$^{39}\text{Ar}_{\text{K}}$ (% cumulated)	$^{40}\text{Ar}_{\text{radiog}}/^{39}\text{Ar}_{\text{K}}$	Age (kyr)
BRG/81/23*					
400	90.4	0.9331	2.4	0.3550	231 ± 34
600	67.0	0.4741	15.3	0.3432	223 ± 7
700	42.8	0.2499	48.8	0.3630	236 ± 5
830	53.1	0.3126	67.0	0.3570	232 ± 6
920	76.2	0.6822	72.8	0.3887	253 ± 17
1,080	87.6	2.690	75.5	0.3431	223 ± 39
1,280	88.5	5.891	100.0	0.3320	216 ± 6
BRG/81/26†					
450	99.2	2.177	1.6	0.3187	205 ± 162
600	95.4	1.005	15.9	0.4544	292 ± 21
700	85.8	0.885	46.3	0.4578	294 ± 11
800	75.7	1.264	61.5	0.7831	503 ± 21
880	77.9	1.770	69.1	1.003	645 ± 45
1,000	84.3	3.958	71.8	0.9426	606 ± 76
1,200	79.9	10.22	90.3	0.9953	639 ± 10
Fuse	83.1	19.35	100.0	1.211	778 ± 22

All are errors quoted at 1σ level.

* Standard ($^{40}\text{Ar}/^{39}\text{Ar}_{\text{K}}$) = 26.82 ± 0.09 ; integrated age = 229 ± 3 ; mass = 10.800 g.

† Standard ($^{40}\text{Ar}/^{39}\text{Ar}_{\text{K}}$) = 27.18 ± 0.06 ; integrated age = 470 ± 8 ; mass = 10.570 g. Standard age¹³ = 17.3 ± 0.12 Myr.

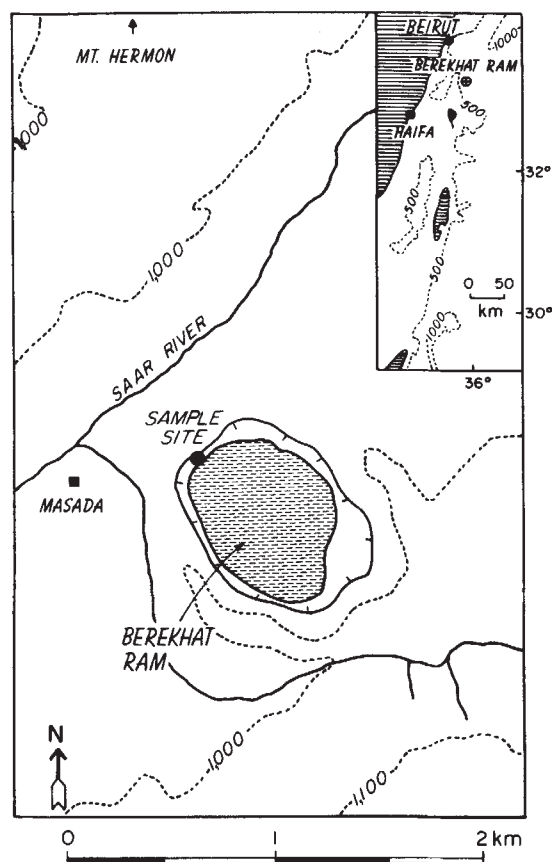


Fig. 1 Location map. Large black dot marks location of archaeological site between two basalt flows on inner rim of crater. Crater is filled by a lake.

The rims of the olivine crystals are altered to a reddish-brown colour; otherwise, the rock appears to be unaltered. The overlying flow is much finer grained, and less vesicular. Fine olivines are set in a mesostasis of very fine plagioclase laths and needles, equant oxide crystals, needles of pyroxene, and basaltic glass.

The upper surface of the lower flow is composed of deeply weathered to spheroidal masses; a sandy red soil fills the interstices. A planar zone, 20 cm above the bedrock, is rich in cobble-to pebble-sized basalt fragments, and contains a dense concentration of artefacts, made of flint and coated with reddish-brown patina. Overlying the artefact-rich layer is a sequence of reddish-brown sandy soils containing minor amounts of rock fragments derived from outcrops 1 km to the north-west on the slopes of Mt Hermon. Today the valley of the Saar River blocks the possible transport of sediment from Mt Hermon to this site. The overlying basalt has baked the uppermost palaeosol and heat-fractured some flint artefacts embedded in it.

Excavation of $\sim 20 \text{ m}^2$ of the archaeological horizon has yielded more than 7,500 lithic artefacts including both tools and waste products. The tools include side scrapers, end scrapers, notches, denticulates, burins, awls and bifaces. The use of both Levallois and discoidal techniques can be seen on the tools as well as on the waste products (flakes, blades, points, and various types of cores). This lithic assemblage is assigned to the Upper Acheulian and resembles that found at other open air sites in the Northern Golan Heights², the Hula Valley³ and Syria⁴.

In an attempt to provide upper and lower bounds to the age of these artefacts, we have carried out conventional K-Ar and $^{40}\text{Ar}/^{39}\text{Ar}$ stepwise heating age determinations on the basalts. For each $^{40}\text{Ar}/^{39}\text{Ar}$ analysis, 10 g of whole rock were neutron-irradiated in the McMaster University reactor along with three samples of the Bern 4B biotite standard. The basalts were step-heated in a radio-frequency furnace while the biotite standards were fused in one step using the University of Toronto continuous argon-ion laser probe⁵.

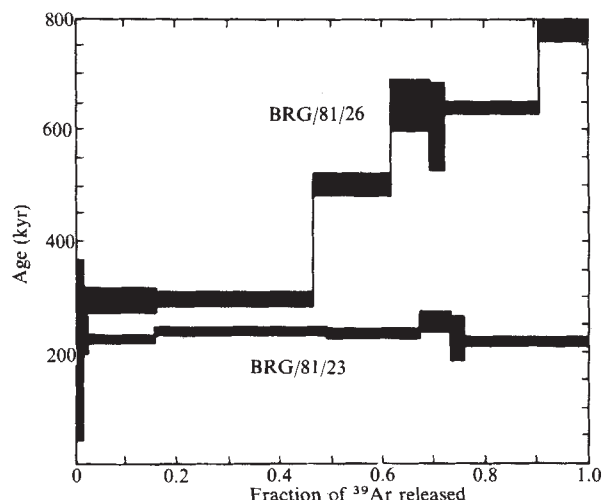


Fig. 2 $^{40}\text{Ar}/^{39}\text{Ar}$ age spectra obtained from lower flow (BRG/81/26) and upper flow (BRG/81/23). Widths of boxes represent $\pm 1\sigma$ errors.

The conventional K-Ar method (Table 1) gave ages of 252 ± 14 (1σ) kyr, and 510 ± 32 (1σ) kyr for the upper (BRG/81/23) and lower (BRG/81/26) flows respectively. The $^{40}\text{Ar}/^{39}\text{Ar}$ analyses (Table 2) essentially agree with the K-Ar age of the upper basalt but show that the lower flow has been disturbed and cannot at present be assigned a reliable age.

Figure 2 shows that the $^{40}\text{Ar}/^{39}\text{Ar}$ age spectrum of the upper basalt exhibits a very well-defined plateau at 233 ± 3 (1σ) kyr for the first 75% of ^{39}Ar released. The corresponding integrated $^{40}\text{Ar}/^{39}\text{Ar}$ age is 229 ± 3 (1σ) kyr, which agrees within 2σ errors with the earlier conventional K-Ar result. Because of the remarkable flatness of the plateau, and the high precision obtained in the run due to the separation of radiogenic and atmospheric argon in the step-heating, we consider 233 ± 3 (1σ) kyr to be a good estimate of the age of this flow.

In contrast, the lower flow yields a very disturbed age spectrum (Fig. 2) with ages increasing from 205 to 780 kyr as the release temperature increases. The integrated age of 470 ± 8 kyr, however, is in good agreement with the conventional K-Ar age. It is impossible to give a single, convincing interpretation of this age spectrum. It seems to exhibit two poorly defined apparent plateaus, one at low temperatures (450–700 °C), another at higher temperatures (880–1,200 °C). The first one, representing 46% of the total ^{39}Ar released, gives an age of ~ 293 kyr, while the second, representing only 29% of the ^{39}Ar released, yields an age of ~ 639 kyr. Several phenomena could be responsible for this, but the most likely possibilities are either argon loss from the lava due to some kind of sample alteration, or the presence of excess argon in this flow. The first alternative would imply that the lower volcanic was probably at least 800 kyr old. The second possibility, however, would suggest a date of about 290 kyr. There is no independent support for either hypothesis. Thus, this lower flow shows, in thin section, only very slight indications of alteration, while the age spectrum does not exhibit the saddle-shape frequently associated with the presence of excess argon⁶. We can only conclude that the lower lava probably crystallized somewhere between 290 kyr and 800 kyr ago. The data from the lower flow emphasize the necessity of using the $^{40}\text{Ar}/^{39}\text{Ar}$ step-heating approach when such extremely young material is being dated. Without observing the two greatly contrasting $^{40}\text{Ar}/^{39}\text{Ar}$ spectra, we would have had no analytical reason to assign more reliability to the one conventional K-Ar age than the other.

These results give a well-defined younger limit to the age of the Acheulian industry at Berekhat Ram of 233 ± 3 kyr. The palaeosol appears to be a cumulus soil⁷ formed by slow accretion by wind and/or water, and may have taken a considerable length of time to accumulate. If it accumulated at a fairly constant rate, then the site would have been occupied close to

the age of the underlying basalt, here estimated to lie between 290 and 800 kyr.

The nearby site of Maayan Barukh also contains a late Acheulian industry. A U-series analysis⁸ of the travertine enclosing this site suggested its age was > 350 kyr. Also, an Upper Acheulian industry at Tabun Cave, Mt Carmel, north-west Israel is inferred⁹ to have been deposited during the last interglacial, ~ 125 kyr. Our present results show that the Upper Acheulian industry of the Levant was in use starting at least 240 kyr BP. Furthermore, use of this industry may have persisted over the subsequent 140 kyr until it was abruptly replaced by the Middle Palaeolithic industries^{10,11}.

We acknowledge the assistance of P. Goldberg in the interpretation of the palaeosols. The potassium analysis was done by P. Kuybida. For financial assistance, we thank NSERC and SSHRC (Canada), the department of antiquities (Israel) and the L.S.B. Leakey Foundation. G.F. was a NATO fellow. We thank O. Bar Yosef and D. Yaalon for their comments.

Received 2 March; accepted 17 May 1983.

1. Mor, D. *Proc. Israel geol. Soc. A. Meet.* 30–32 (1981).
2. Goren, N. *Colloq. Int. Cent. natn. Rech. Scient.* 598, 193–205 (1981).
3. Ronen, A., Ohel, M., Lamdan, M. & Assaf, A. *Israel Explor. J.* 30, 17–33 (1980).
4. Goren, N. *Quartär* 29/30, 105–131 (1979).
5. York, D., Hall, C. M., Yanase, Y., Hanes, J. A. & Kenyon, W. J. *Geophys. Res. Lett.* 8, 1136–1138 (1981).
6. Lanphere, M. & Dalrymple, G. B. *Earth planet. Sci. Lett.* 32, 141–148 (1976).
7. Yaalon, D. & Ganor, E. *Soil Sci.* 116, 146 (1973).
8. Schwarcz, H. P., Goldberg, P. & Blackwell, B. *Israel J. Earth-Sci.* 29, 157–165 (1980).
9. Jelinek, A. J. *Science* 216, 1369–1375 (1982).
10. Bar-Yosef, O. & Goren, N. *Un. int. Sci. Preh. Protohist. X, Commun.* 6, 28–38 (1981).
11. Bar-Yosef, O. A. *Rev. Anthropol.* 9, 101–133 (1980).
12. Steiger, R. & Jäger, E. *Earth planet. Sci. Lett.* 36, 359–362 (1977).
13. Hall, C. M., Walter, R. C., Westgate, J. A. & York, D. Preprint (University of Toronto, 1983).

Separate orientation and releaser components in a sex pheromone

J. W. S. Bradshaw, R. Baker & J. C. Lisk

Chemical Entomology Unit, Departments of Biology and Chemistry, Building 3, The University, Southampton SO9 5NH, UK

The sex pheromones, with which female moths attract males, normally consist of a mixture of several chemicals, usually aliphatic alcohols, acetates and aldehydes¹. In the small number of species which have been studied in detail, particular combinations of compounds have been shown to elicit each of the major stages of the approach of the male to the female^{2–4}. However, little distinction has been made between the chemical signals which allow the male to orientate towards the female, and those which elicit specific responses such as landing and courtship behaviour. Recent studies on the mechanisms whereby male moths navigate towards pheromone-producing females have concentrated on single chemicals, or fixed blends of chemicals, and have not considered the separate roles of individual compounds within a blend^{5,6}. In our studies of the responses of male pine beauty moths, *Panolis flammea* (Schiff.) to the sex pheromone produced by the female, we have now demonstrated that both landing and copulatory behaviour are elicited by different chemical stimuli to those which serve to orientate the male at each of these stages.

The female pine beauty moth produces a blend of three compounds⁷: (Z)-9-tetradecenyl acetate (Z-9:14Ac), (Z)-11-tetradecenyl acetate (Z-11:14Ac) and (Z)-11-hexadecenyl acetate (Z-11:16Ac) in the proportions 100:1:5. In field experiments using sticky traps baited with synthetic samples, Z-9:14Ac on its own caught males, but neither of the Z-11 compounds, nor their binary mixture, attracted males. The binary mixture of Z-9:14Ac and Z-11:14Ac was more effective than Z-9:14Ac alone, and was equally as effective as the third binary mixture, Z-9:14Ac and Z-11:16Ac, and the three-component combination (see Table 1).

In a separate trapping experiment, the three-component blend was as effective as any of the 24 other combinations which can be produced by setting the levels of Z-11:14Ac at 0.1, 0.3, 1, 3 and 10% and Z-11:16Ac at 0.5, 1.5, 5, 15 and 50% of Z-9:14Ac. Thus, the long-range attractant components of the *P. flammea* sex pheromone are apparently Z-9:14Ac plus either one of the two Z-11 compounds. This raised some doubt as to whether both Z-11 compounds are essential components of the sex pheromone, therefore the behaviour of male *P. flammea* to the three most effective combinations was studied in a laboratory wind tunnel⁸.

In a typical experiment, five male moths were released into the tunnel approximately 1.5 m downwind of an elevated platform. After 2 min a lure was placed on the platform, producing a plume of odour extending downwind. For the next 5 min, any moth flying upwind across a horizontal line 50 cm long that was 50 cm downwind of the lure, was recorded on videotape, and

Table 1 Total numbers of male *P. flammea* caught in sticky traps, in three field experiments

Lure	Total catch
Z-9:14Ac (500 µg)	89*
Z-11:14Ac (5 µg)	0†
Z-11:16Ac (40 µg)	0†
Z-11:14Ac+Z-11:16Ac (5+40 µg)	0†
Z-9:14Ac (100 µg)	58‡
Z-9:14Ac+Z-11:14Ac (100+1 µg)	184§
Z-9:14Ac+Z-11:14Ac (500+5 µg)	43
Z-9:14Ac+Z-11:16Ac (500+40 µg)	72
Z-9:14Ac+Z-11:14Ac+Z-11:16Ac (500+5+40 µg)	65

Synthetic samples of the three compounds (>99% pure; abbreviations defined in the text) were loaded on to red rubber sleeve stoppers (West Pharmarubber, UK) placed centrally in the sticky area of large Delta traps (PLA, Insect Monitoring, Southampton, UK). Z-11:16Ac was loaded at 8%, rather than 5%, to compensate for its reduced release rate from this substrate, compared with the tetradecenyl acetates. The traps were placed on wooden posts, 1.5 m high, 25 m apart in forest rides. The experiments were set out in Ringwood Forest, Hampshire (first and third parts of the table) and Bareagle Forest, Dumfries and Galloway (second part of the table), in areas planted with *Pinus contorta*, during April 1981 and April 1980, respectively. There were five replicates of each treatment, arranged in randomized complete blocks.

*†‡§|| Numbers within a section indicated by the same symbol are not significantly different by analysis of variance ($\sqrt{x+1}$ transformation) and Newman and Keuls' multiple-range test ($P<0.05$)⁹.

the tracks of the moths were analysed for landing on the platform, and for attempted copulation with the source of the chemicals (defined as contacts with the lure, with genitalia extended, lasting ≥ 1 s). Landing was observed frequently with treatments including Z-11:14Ac, but only once for the treatment excluding this compound, thus it must be important for landing (Table 2A). However, all three components elicit the full sequence of flight-landing-copulation much more frequently than either binary mixture. Similar experiments using Z-9:14Ac alone as lure did not yield any upwind flights from more than 50 males.

The roles of Z-9:14Ac and Z-11:14Ac in landing behaviour were investigated by placing each in turn between the mesh at the upwind end of the tunnel, and the fan. This distributed the odour throughout the flight chamber, and produced a contrasting stimulus to that received from the remaining two components, which were placed on the platform in the flight chamber, as before. The dose of the upwind compound was increased five times to compensate for the increased dilution caused by this method of placement. When Z-9:14Ac was presented in diffused form, the moths became very active and flew very fast upwind and downwind. They appeared unable to locate the 'plume' containing the two Z-11 compounds. However, when Z-11:14Ac was presented in diffused form, the moths were able to locate the plume and complete their normal sequence of behaviour (Table 2B). This indicates that while Z-11:14Ac 'switches on' landing, it is the plume of Z-9:14Ac which orientates this response.

These two experiments do not indicate the exact stimuli releasing copulation, as copulation follows landing, which is infrequent in the absence of Z-11:14Ac. Thus, single males were confined for 5 min each with seven different combinations of the three compounds (Table 3). The combination of Z-9:14Ac and Z-11:16Ac was presented both as a single source, and as two sources, each containing one compound. With the requirement for landing removed, the combination of Z-9:14Ac and Z-11:16Ac elicited attempted copulation, but neither Z-9:14Ac nor Z-11:16Ac did so alone, and when present as two distinct sources, copulation was directed to Z-9:14Ac. There is no evidence for any dominant role of Z-11:14Ac in this phase of the behaviour.

From these experiments, we propose a hierarchy of stimuli which elicit various stages in the behaviour of a male pine beauty moth approaching a female. Long-range attraction is mediated by Z-9:14Ac in combination with any one, or both, of Z-11:14Ac and Z-11:16Ac. Captures of moths in the field with Z-9:14Ac alone may arise due to a situation similar

Table 2 Responses of male *P. flammea* in a flight tunnel to combinations of sex pheromone components

		No. of males tested	No. of tracks	% Landing	Landed, % copulating
A All compounds released together					
Z-9:14Ac (500 µg)+Z-11:14Ac (5 µg)		30	8	88*	14‡
Z-9:14Ac (500 µg)+Z-11:16Ac (40 µg)		30	14	7†	(0)
Z-9:14Ac (500 µg)+Z-11:14Ac (5 µg)+Z-11:16Ac (40 µg)		30	16	81*	77§
B Compounds released separately					
Diffuse	Point source				
Z-9:14Ac (2.5 mg)	Z-11:14Ac (5 µg)+Z-11:16Ac (40 µg)	15	>20	0	—
Z-11:14Ac (25 µg)	Z-9:14Ac (500 µg)+Z-11:16Ac (40 µg)	15	11	91*	70§

Compounds were formulated as in Table 1 legend. The walls of the flight tunnel consisted of thin clear plastic tubing which, when inflated, forms a cylindrical tunnel of diameter 75 cm. One end of the tubing was attached to the housing of a fan; 40 cm downwind of the fan the tubing was constricted around a hoop-shaped frame, 55 cm in diameter and containing disks of aluminium honeycomb (1 cm diameter mesh), and fine stainless steel mesh. A second frame, without the aluminium honeycomb, was attached to the tubing downwind of the first to delineate a flight chamber 2 m long. Exhaust air was ducted out of the building. To measure the approximate distribution of volatile chemicals within the tunnel, sources of ammonium chloride 'smoke' were placed (a) on a platform 18 cm high, 30 cm downwind of the upwind frame and (b) on an identical platform equidistant between the upwind frame and the fan. At a, a 'plume' was formed, approximately cone-shaped with the source at the tip, and the base of ~50 cm diameter at the downwind mesh. At b, the smoke was distributed throughout the flight chamber. Lures were placed on identical platforms on 5 inch×3 inch index cards, which were used only once each. Tests with insects were performed under dim red light in the first 90 min of scotophase, at 15–18 °C, wind speed 0.17 m s⁻¹.

*†‡§|| Mean values indicated by the same symbol are not statistically significant by χ^2 tests ($P<0.01$).

|| Tracks were often too fast to record accurately from video tape.

Table 3 Copulatory behaviour of males confined with various combinations of sex pheromone components, in two experiments

Lure	No. of males tested	No. of males attempting copulation
A Blank	12	0*
Z-9:14Ac	12	0*
Z-9:14Ac+Z-11:14Ac+Z-11:16Ac	12	9†
B Blank	8	0‡
Z-9:14Ac+Z-11:14Ac	8	0‡
Z-11:16Ac	8	0‡
Z-11:14Ac+Z-11:16Ac	8	1‡
Z-9:14Ac+Z-11:16Ac	8	7§
Z-9:14Ac+Z-11:14Ac+Z-11:16Ac	8	5§
Z-9:14Ac with Z-11:16Ac	8	6§
Z-11:16Ac with Z-9:14Ac		1‡

Individual male *P. flammea* were confined for 5 min in clear plastic boxes, 150×90×60 mm, containing compounds formulated as indicated in Table 2A, and placed centrally on the base of the box. The last two lines of the table refer to the same treatment, where two lures were used, one containing Z-9:14Ac and one Z-11:16Ac, placed 30 mm apart. Copulatory attempts with each source are recorded separately.

*†‡§ Values within A or B indicated by the same symbol are not statistically significant by χ^2 tests ($P < 0.05$).

to that in the second wind tunnel experiment, where an undirected source of Z-11:14Ac was provided by nearby lures containing this compound, or by nearby 'calling' females. As the level of Z-11:14Ac increases near to its source, it reaches a threshold which initiates a search for a landing site, guided by the plume of Z-9:14Ac, and presumably, visual and tactile stimuli. After landing, the male approaches the female, guided by Z-9:14Ac and probably visual stimuli, and copulation is initiated by a high concentration of Z-11:16Ac and Z-9:14Ac. Thus, it seems that, while the males retain Z-9:14Ac as an orientation cue throughout, their response is modified by suitably high concentrations of the Z-11 compounds. There are almost certainly other stages not included in this scheme, including changes of velocity and flight pattern in the last few metres of approach, as observed for another noctuid, *Spodoptera littoralis*⁵.

In most studies of behaviour in response to multicomponent sex pheromones, the implicit assumption has been made that male moths react to a given ratio of two or more compounds as if they were a single signal. We have demonstrated that, at least in the case of *P. flammea*, the individual compounds, which in combination mediate two of the major stages of sex attraction, provide separate cues for relaying the 'switch' to the next stage of behaviour, and orientating that stage. The information from these cues is presumably integrated in the central nervous system, to produce the normal behaviour pattern, and it is only when conflicting signals are presented, as in our experiments, that the orientation cue can be identified. Precise knowledge of the roles of each of the pheromone components is therefore essential to any study of orientation mechanisms.

We thank J. T. Stoakley and D. Bevan (Forestry Commission, UK) for the supply of insect pupae and for arranging field sites, R. Mothershaw for technical assistance, and M. C. Birch and O. T. Jones for review of the manuscript. This work was supported by grant GR/B08561 from the UK SRC.

Two types of sex determination in a nematode

Jonathan Hodgkin

MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK

Sex in the nematode *Caenorhabditis elegans* is normally determined by a genic balance mechanism¹, the ratio of X chromosomes to autosomes, so that XX animals are self-fertilizing hermaphrodites and XO animals are males^{2,3}. However, recessive mutations of the autosomal gene *tra-1 III* cause both XX and XO animals to develop into males⁴, and a linked dominant mutation causes both XX and XO animals to develop into females⁵. Here I show that these two kinds of mutation are allelic, and that stable mutant strains can be constructed in which sex is determined not by X-chromosome dosage but by the presence or absence of a single active gene. In these strains the autosomes carrying the *tra-1* locus are in effect homomorphic Z and W sex chromosomes, and the sexes are homogametic ZZ males and heterogametic ZW females, in contrast to the wild-type arrangement of homogametic XX hermaphrodites and heterogametic XO males.

Males and hermaphrodites of *C. elegans* differ extensively in structure and post-embryonic development⁶. Males have a one-armed testicular gonad which opens posteriorly, and specialized tail structures used in copulation (Fig. 1). Hermaphrodites have a two-armed gonad which opens at a vulva located in the middle of the body, and no specialized copulatory structures. Hermaphrodites and females differ only in the differentiation of gametes: hermaphrodite gonads make both sperm and oocytes, while female gonads make only oocytes. Females do not occur naturally in strains of *C. elegans*.

The primary mechanism of sex determination in wild-type *C. elegans* is similar to that in *Drosophila*⁷: a high ratio of X chromosomes to autosomes is hermaphrodite-determining, while a low ratio is male-determining. Also as in *Drosophila*, there are multiple sites on the X chromosome that contribute additively to the X:autosome ratio³. The sex-determining effect of this ratio can be circumvented by mutations in a small number of autosomal genes, the *tra*, *her* and *isx* genes^{4,5,8}. The most extreme effects are caused by mutations in *tra-1 III*.

Recessive mutations of *tra-1* have been described previously⁴: they have no effect on XO animals, which develop into normal fertile males, and they cause XX animals also to develop into phenotypic males. Several null alleles, such as *e1099*, have been identified, which cause perfect masculinization of all non-gonadal characters in XX animals. The gonads are also masculinized in these XX males, but gonadal development is variable and frequently abnormal. Nevertheless, about one-third of the animals are fertile XX males that can sire progeny. These mutations are regarded as null alleles first because in complementation tests they behave similarly to a large genetic deficiency (*eDf2*) that includes the *tra-1* locus⁵, and second because one allele of this type (*e1781*) is suppressed by *sup-7* X, which is now known to be an amber suppressor^{9,10}.

A dominant allele of *tra-1*, *e1575*, has effects opposite to those of the recessive mutations: both XO and XX animals are feminized by this mutation⁵. XX and XO animals heterozygous for the dominant allele (*e1575/+*) have hermaphrodite anatomy, but usually make no sperm, that is, they are functional females rather than hermaphrodites. A minority of the heterozygotes make a few sperm, and are self-fertile, while homozygotes of both XX and XO types are invariably females. Genetic mapping showed that *e1575* is tightly linked to the *tra-1* locus (recombination <0.5%). The deficiency *eDf2* includes all of this region, but has no dominant effect on sexual

Received 6 December 1982; accepted 7 June 1983.

- Roelofs, W. L. & Carde, R. T. *Rev. Ent.* **22**, 377-405 (1977).
- Baker, T. C., Carde, R. T. & Roelofs, W. L. *J. chem. Ecol.* **2**, 333-352 (1976).
- Baker, T. C. & Carde, R. T. *Envir. Ent.* **8**, 956-968 (1979).
- Linn, C. E. & Gaston, L. K. *Envir. Ent.* **10**, 751-755 (1981).
- Murlis, J. & Bettany, B. W. *Nature* **268**, 433-434 (1977).
- Kennedy, J. S., Ludlow, A. R. & Sanders, C. J. *Nature* **288**, 475-477 (1980).
- Baker, R., Bradshaw, J. W. S. & Speed, W. *Experientia* **38**, 233-234 (1982).
- Jones, O. T., Lomer, R. A. & Howse, P. E. *Physiol. Ent.* **6**, 175-181 (1981).
- Snedecor, G. W. & Cochran, W. G. *Statistical Methods* (Iowa State University Press, 1967).

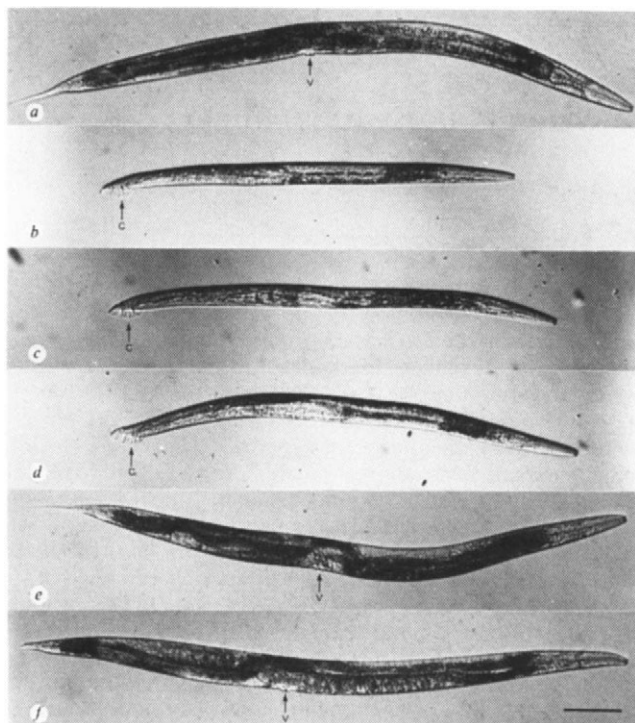


Fig. 1 Light micrographs of young adult *C. elegans* nematodes. *a*, Wild-type XX (hermaphrodite); *b*, wild-type XO (male); *c*, *tra-1*(null/null) XX (male); *d*, *tra-1*(null/null) XO (male); *e*, *tra-1*(null/dom) XX (female); *f*, *tra-1*(null/dom) XO (female). Alleles: *tra-1*(null) (null allele) = *e1099* and *tra-1*(dom) (dominant allele) = *e1575*. Arrows indicate centrally located vulva (v) in *a*, *e* and *f*; specialized male tail copulatory structures (c) in *b*, *c* and *d*. Scale bar, 100 µm.

phenotype, so *e1575* must be a neomorphic alteration of some kind. These facts, and the observation that it has dominant effects opposite to those of the recessive null mutations, made it seem likely that *e1575* causes constitutive expression of the *tra-1* gene. If this is so, it should be possible to revert the dominant effects of *e1575* by introducing a null mutation of *tra-1* in *cis* to *e1575*. This was achieved by selecting for self-fertility in *e1575*/+ XX animals. Many null mutations were obtained by this technique which were indistinguishable from the original set of null alleles such as *e1099* or *e1781*. Some of these new null mutations are amber-suppressible, so that in the presence of *sup-7* the double mutant is converted from a recessive masculinizing allele back into a dominant feminizing allele. Attempts to separate one of these double mutants have been unsuccessful, indicating that the amber mutation is tightly linked to the dominant mutation (recombination <0.1%). These results are consistent with the belief that *e1575* is a *cis*-dominant constitutive mutation of *tra-1*.

The heterozygote of the constitutive allele and a null allele (for example, *tra-1*(*e1575*/*e1099*)) is female, like homozygous *e1575*. If such a female is fertilized by a homozygous *tra-1*(*e1099*) male, progeny of the same genotypes as the parents are produced, as shown in Fig. 2. This establishes a male/female (gonochoristic) strain in which the chromosome (autosome III) carrying *e1575*(*tra-1*(dom)) is a W sex chromosome and the chromosome carrying a null (*tra-1*(null)) is a Z sex chromosome. ZZ animals are male and ZW animals are female, regardless of X-chromosome dosage (Fig. 1). If the null allele is one of the original set, such as *e1099*, the strain may eventually revert back to self-fertility by recombination between *e1575* and the null, and indeed, rare self-fertile animals have been recovered from the heterozygotes with several different null alleles. If the null is a double mutant that already carries *e1575*, this cannot occur, and the strain is stable. These ZZ/ZW male/female strains grow very poorly compared with the wild-type hermaphrodite strain, because they depend on mating for

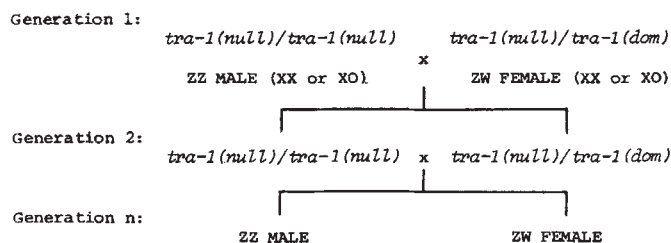


Fig. 2 An alternative sex determination scheme.

propagation and because the fertility of *tra-1*(null) XX males is low. However, they are viable and can be maintained for many generations, thus demonstrating the possibility of single locus sex determination in *C. elegans*.

The observation that a constitutive *tra-1* mutation causes *C. elegans* to develop as a female may be significant in evolutionary terms. Most nematode species have male and female sexes^{1,11} and hermaphroditism is therefore likely to be a secondary specialization, as originally suggested by Maupas¹². The ZZ/ZW strain could therefore represent a return to the ancestral condition. Also, since the wild-type sexes are male and hermaphrodite (not female), there must have evolved a mechanism for controlling *tra-1* activity in both wild-type sexes. The other major sex-determining genes (*her-1*, *tra-2*, *tra-3* and probably *isx-1*) may provide part of this mechanism, because they appear to act at a level between the X:autosome genic balance and the crucial gene *tra-1* (ref. 5).

Irrespective of these speculations, it is clear that all aspects of the sexual phenotype of *C. elegans* can be controlled by the state of a single locus, *tra-1*, which is therefore comparable to the *M* loci of certain flies¹³ and mosquitoes¹⁴, or the putative structural gene for the H-Y antigen of mammals¹⁵, as a single gene switch for sex determination. Simple switch mechanisms of this type may underlie sex determination in many other organisms.

Received 11 April; accepted 9 June 1983.

- White, M. J. D. *Animal Cytology and Evolution* (Cambridge University Press, 1973).
- Brenner, S. *Genetics* **77**, 71-94 (1974).
- Madl, J. E. & Herman, R. K. *Genetics* **93**, 393-402 (1979).
- Hodgkin, J. A. & Brenner, S. *Genetics* **86**, 275-287 (1977).
- Hodgkin, J. *Genetics* **96**, 649-664 (1980).
- Sulston, J. E. & Horvitz, H. R. *Dev. Biol.* **56**, 110-156 (1977).
- Lauge, G. in *Genetics and Biology of Drosophila* Vol. 2d (Academic, New York, 1977).
- Nelson, G. A., Lew, K. K. & Ward, S. *Dev. Biol.* **66**, 386-409 (1978).
- Waterston, R. H. *Genetics* **97**, 307-325 (1981).
- Wills, N. et al. *Cell* **33**, 575-583 (1983).
- Triantaphyllou, A. C. & Hirschmann, H. A. *Rev. Phytopath.* **2**, 57-80 (1964).
- Maupas, E. *Archs Zool. exp. gen.* **8**, 463-624 (1900).
- Mainx, F. *Am. Nat.* **98**, 415-430 (1964).
- Baker, R. H. & Sakai, R. K. *J. Hered.* **67**, 289-294 (1976).
- Wachtel, S. S., Ohno, S., Koo, G. C. & Boyse, E. A. *Nature* **257**, 235-236 (1975).

Single-channel currents activated by light in *Limulus* ventral photoreceptors

Juan Bacigalupo & John E. Lisman

Department of Biology, Brandeis University, Waltham, Massachusetts 02254, USA

Sensory stimuli alter the membrane conductance of receptor cells. Noise analysis studies in several types of receptors¹⁻⁴ suggest that ionic channels mediate this change in conductance, but no direct evidence for such channels has been obtained. We have now investigated the mechanism of ion permeation underlying the light-induced conductance in the ventral photoreceptors of *Limulus polyphemus* by means of the patch-clamp technique⁵ and have resolved single-channel currents that are activated by light.

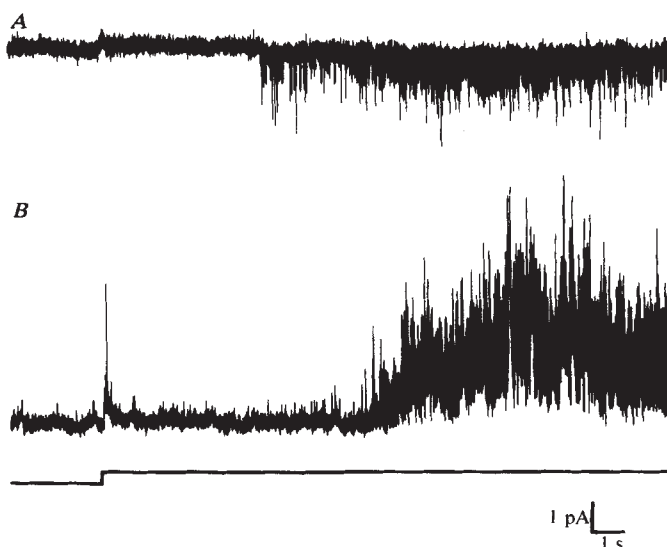


Fig. 1 Effect of light and voltage on *Limulus* single-channel activation. A light stimulus of log intensity -3.0 (see Fig. 2) was applied to the cell. The trace at the bottom of the figure indicates the timing of the stimulus. A, Current recorded from the patch at ~ -45 mV (resting potential). B, Current recorded from the patch depolarized to $\sim +55$ mV. The initial outward current seen in both traces shortly after the onset of the light stimulus was due to the charging of the patch membrane capacitance during the rise of the receptor potential. The pipette was connected to a current-to-voltage transducer (OPA 101, feedback resistor = $10^9 \Omega$) and the recordings filtered with a 4-pole Bessel filter. Bandwidth 500 Hz. For the purpose of sonicating the preparation (80 V for 2–3 min; sonicator model G112SP1T from Lab. Supply Co. Inc., Hicksville, New York), the chamber containing the nerve pretreated with Pronase was wrapped in parafilm and suspended in the sonicator. Patch pipette was filled with artificial seawater of the following composition (in mM): 425 NaCl, 10 KCl, 10 CaCl_2 , 22 MgCl_2 , 26 MgSO_4 , 15 Tris-HCl, pH 7.8. Temperature 21°C .

Ventral nerves that contain the photoreceptors were desheathed and treated with 2% Pronase (Calbiochem) for ~ 1 min in the standard way⁶. Patch pipettes had tip diameters of $\sim 0.5 \mu\text{m}$ and were filled with artificial seawater (for composition see Fig. 1 legend). The glial cells and connective tissue surrounding each photoreceptor were removed as described by Stern *et al.*⁷, exposing its plasma membrane. Ventral photoreceptors have two kinds of lobes^{7,8}, one of which, the A-lobe, has a flat and unfolded membrane^{7,8}. We were able easily to obtain seal resistances of 10^9 – $10^{11} \Omega$ between the patch pipette and the membrane of the A-lobe, but observed only voltage-dependent channels. The R-lobe, which is the light-sensitive part of the cell, bears microvilli on its surface⁷. Without further treatment of the cells, we obtained seals in this lobe only with difficulty (requiring high suction). The recordings had a sporadic uninterpretable noise and in no case showed light-activated channels. To obtain better recordings, we briefly placed the preparation in a sonicator. Cells treated in this way appeared unchanged under the light microscope and had normal resting potentials (56.6 ± 12.7 mV in 14 cells), normal receptor potentials and normal responses to single photons (quantum 'bumps'). With such cells we obtained seals more easily ($\sim 30\%$ of the attempts were successful) and of 77 membrane patches, 4 contained light-activated channels.

Figure 1 shows the activation of single-channel currents by light. The channel activity returns to the dark level after the light is turned off (not shown). These events are recognized as single-channel currents if the time scale of the recordings is expanded (see Fig. 3A). The single-channel currents are inward at the resting potential, as is the macroscopic light-dependent current. As ventral photoreceptors depolarize in response to light, it was important to determine whether the channels were activated by depolarization rather than by a more direct effect

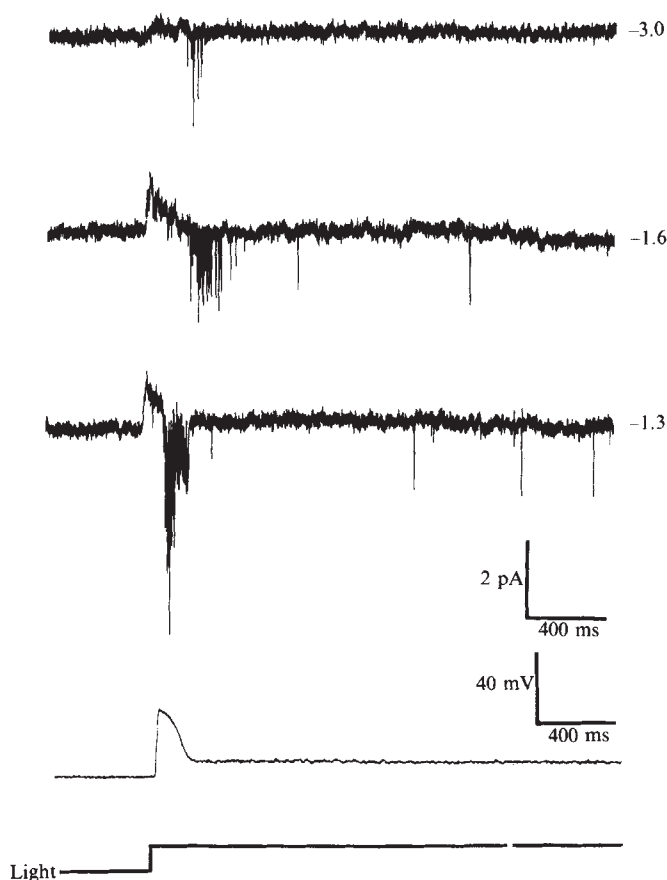


Fig. 2 Patch-clamp currents in response to light stimulus of different intensities. The numbers on the right denote the relative light intensities ($\log I/I_0$, where I_0 , the unattenuated 520 nm light intensity, is 1.4×10^{14} photons $\text{cm}^{-2} \text{s}^{-1}$). The bottom trace is the receptor potential recorded simultaneously with the lowest current trace (log intensity -1.3). Resting potential -47 mV. The step indicates the timing of the stimulus. In this patch, the initial outward currents have a contribution due to voltage-dependent channels in addition to the capacitive current (Fig. 1). Bandwidth 500 Hz, temperature 18°C .

of light. The initial part of Fig. 1B shows that channels were not activated in the dark on depolarization of the membrane patch to $+55$ mV (resting potential -45 mV). However, when a light stimulus was applied while the patch remained depolarized, a dramatic increase in channel activity occurred. The light-activated current was now outward because $+55$ mV is more positive than the zero-current potential of the channel (see below). This indicates that light, and not the depolarization induced by light, caused activation of the channels.

The frequency of channel openings depended on light intensity and time in a manner qualitatively similar to the macroscopic current. In response to dim light, the macroscopic current rises after a 150-ms latency to a small maintained value⁹; such a response at the single-channel level is illustrated in Fig. 1. In response to brighter light, the macroscopic response rises after a shorter latency to a higher value, then falls to a lower maintained level (bottom trace of Fig. 2; ref. 9). Similarly, the single-channel activity increased with light intensity (Fig. 2) and showed a high initial level of activity followed by a much lower maintained value (Figs 2, 3). The latency of single-channel activity was intensity-dependent, but was always longer by a few hundred milliseconds (Fig. 2) to a few seconds (Fig. 1) than that of the macroscopic current. This unusually long latency may be due to an abnormality produced by local deformation of the cell by the suction pipette. Alternatively, we may be preferentially sampling a small subpopulation of the light-activated channels that normally open with a latency greater than average.

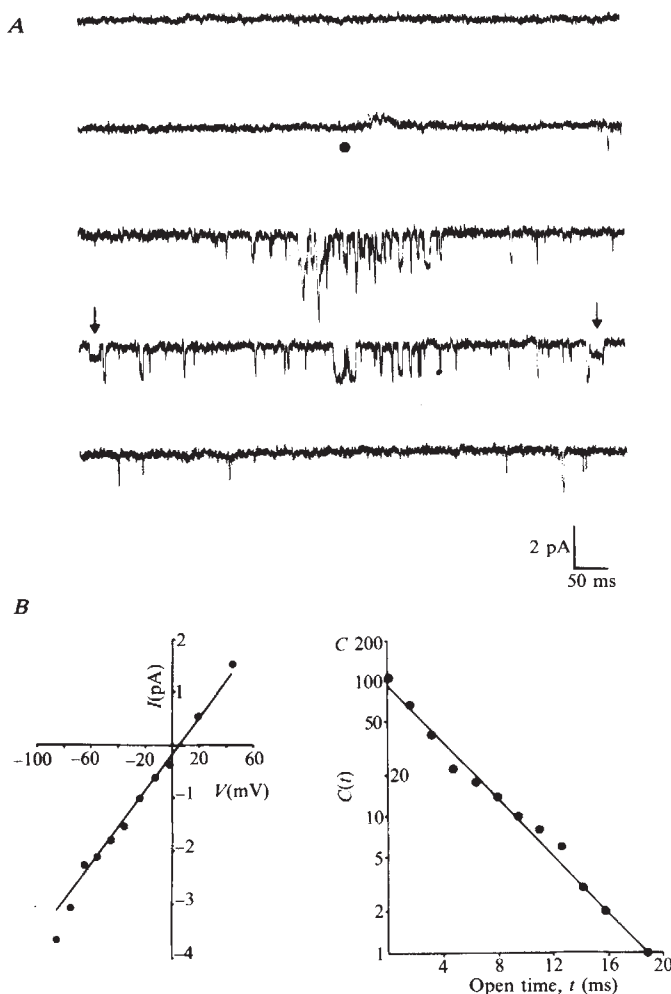


Fig. 3 A, High temporal resolution of single-channel current opening in response to a light stimulus of log intensity -1.6 . Single-channel currents of amplitude 1.5 pA are seen, as well as a few 0.6 -pA single-channel currents (arrows). The dot indicates the onset of the light stimulus. The limiting frequency was determined by the τ of the recording system: 0.8 ms. B, Current-voltage (I - V) curve for the channels of the same patch. The line is the best fit to the experimental points within the voltage range -65 to $+20$ mV. The conductance calculated from this plot is 35 pS. Resting potential -40 mV. C, Semilog plot of the cumulative distribution of the open times of the channels, $C(t)$. Each point represents the number of events having an open time $\geq t$. Only those events occurring after the initial burst and with amplitude ≥ 0.8 pA were considered. The curve fitted to the points is an exponential with $\tau = 4.2$ ms. The average of the open times (105 events) is 3.6 ± 0.4 ($\pm$ s.e.m.) ms. Temperature 18°C .

The currents through single light-activated channels are shown at high time resolution in Fig. 3A. In this patch the single-channel current was ~ 1.5 pA. Some of the very brief events look smaller because their open time was shorter than the time constant of our recording system (~ 0.8 ms). The maximal current through the patch as a whole was ~ 4.5 pA, indicating that the patch contained at least three light-activated channels. A current-voltage curve for the open single light-activated channels in this patch is shown in Fig. 3B. The slope conductance, in the range -65 to $+20$ mV, is 35 pS (45 and 50 pS in other experiments). The zero-current potential obtained from Fig. 3B is $\sim +6$ mV, which is the value expected for the macroscopic light-dependent current⁹. Figure 3A shows a small number of single-channel currents of ~ 0.6 pA. We saw these channels in other patches containing light-activated channels, but we do not know whether they represent different channels or a different conductance state of the same channel. Figure 3C shows that the distribution of open times of the

light-activated channels can be fit by an exponential with $\tau = 4.2$ ms, which was the mean open time of the channel in this patch (at 18°C). The mean open times obtained in other patches were 1.8 (18°C), 1.2 (21°C) and 1.8 ms (21°C). In contrast, values for single-channel conductance and mean open time estimated by noise analysis (18 pS and 18.7 ms, respectively; ref. 1) differ from those measured by us, possibly due to the limitations involved in applying noise analysis to photoreceptors¹⁰.

With the values that we obtained for the single-channel conductance and mean open time we can estimate the number of channel openings during a typical quantum bump (2 nA amplitude and 100 ms duration). The average number of electronic charges passing through the membrane during such an event is 6×10^8 , at -50 mV and 18°C (ref. 11). Under the same conditions, 2.4×10^4 electronic charges are associated with the average single-channel current. From these values we can estimate that $\sim 10^4$ channel openings occur during a bump. It is also of interest that the open time of the channels is short compared with the duration of the quantum bump. Thus, the falling phase of the bump ($\tau = 14.4$ ms; ref. 1) is not rate-limited by the closing of the channels, but rather follows the concentration of the molecule that causes the channels to open.

Patch-clamp recordings on vertebrate rods suggest a unitary light-dependent conductance of 50 fs (ref. 12). This value is compatible with a carrier, but does not exclude the possibility of a channel having an uncommonly low conductance. The evidence presented here suggests that, in *Limulus*, ion channels having common characteristics carry the light-activated current and, more generally, that a sensory stimulus can lead to the opening of ion channels in receptor cells.

We thank Peter O'Day, Wolfgang Hanke, Chris Miller, Irwin Levitan, Eve Marder, Marc Goldring, Cecilia Vergara, Phyllis Robinson, Kevin Chinn and Jeff Stern for useful discussions and comments on the manuscript; Dave Eaton for his machining and Irina Burdenko for typing the manuscript. This work was supported by NIH grant EYO1496 and NSF grant BNS 8105746.

Received 18 January; accepted 12 May 1983.

1. Wong, F. *Nature* **276**, 76-79 (1978).
2. Lamb, T. D. & Simon, E. J. *J. Physiol., Lond.* **263**, 257-286 (1976b).
3. Schwartz, E. A. *J. Physiol., Lond.* **272**, 217-246 (1977).
4. De Felice, L. J. & Alkon, D. L. *Nature* **269**, 613-615 (1977).
5. Hamill, O. P., Marty, A., Neher, E., Sakmann, B. & Sigworth, F. S. *Pflügers Arch. ges. Physiol.* **391**, 85-100 (1981).
6. Lisman, J. E. & Sheline, Y. J. *gen. Physiol.* **68**, 487-501 (1976).
7. Stern, J., Chinn, K., Bacigalupo, J. & Lisman, J. E. *J. gen. Physiol.* **80**, 825-838 (1982).
8. Calman, B. G. & Chamberlain, S. C. *J. gen. Physiol.* **80**, 839-862 (1982).
9. Millicchia, R. & Mauro, A. J. *gen. Physiol.* **54**, 331-351 (1969).
10. Fain, G. L. & Lisman, J. E. *Prog. Biophys. molec. Biol.* **37**, 91-147 (1981).
11. Goldring, M. thesis, Brandeis Univ. (1982).
12. Detwiler, P. B., Conner, J. D. & Bodoia, R. D. *Nature* **300**, 59-61 (1982).

Natural killer-like cells found in B-cell compartments of human lymphoid tissues

Diponkar Banerjee & Robert F. Thibert

Immunopathology Laboratory, Department of Pathology, University Hospital, University of Western Ontario, London, Canada N6A 5C1

Natural killer (NK) cells have been implicated in natural resistance to tumours^{1,2}, particularly lymphomas³. Recently, they have been implicated in the regulation of human erythropoiesis⁴. The distribution of NK cells in peripheral lymphoid tissues has not been fully documented. NK activity in human or murine lymph nodes is significantly lower than in the spleen or peripheral blood^{5,6}. In Hodgkin's disease, affected lymph nodes and spleens have higher NK activity than non-involved nodes or spleens⁷. Using a monoclonal antibody (Leu 7) that defines the human NK antigen HNK-1 (ref. 8), only a small

number of HNK-1⁺ cells were found in lymph nodes, tonsils or thymus, with greater numbers in the spleen⁸. We report here the topographical distribution of HNK-1⁺ cells in normal human lymph nodes, tonsils and spleen and preliminary observations in various forms of malignant lymphoma. The most striking observation was that HNK-1⁺ cells were not distributed uniformly in these tissues but specifically localized in the B-cell compartments (follicular centres).

Double-marker studies showed that intrafollicular HNK-1⁺ cells in lymph nodes, tonsils or spleens expressed a pan-T antigen Leu 1, but did not express cytotoxic-suppressor cell antigen Leu 2, inducer-helper cell antigen Leu 3, Ia antigen or myeloid-monocytic antigen M1. Paracortical HNK-1⁺ cells in lymph nodes and tonsils had identical phenotypes to those within follicular centres. On the other hand, scattered HNK-1⁺ cells in splenic red pulp lacked T-cell antigens but expressed M1. The association of HNK-1⁺ cells with follicular centres was preserved in malignant lymphomas of follicular centre origin. In diffuse lymphomas of either B-cell or peripheral T-cell origin, the HNK-1⁺ cells were randomly distributed in variable numbers. In the normal thymus or thymic lymphoma (T-

lymphoblastic lymphoma with convoluted nuclei) only rare HNK-1⁺ cells were identified. In a case of Hodgkin's disease, HNK-1⁺ cells were numerous in the affected node. The observation of specific association of HNK-1⁺ Leu 1⁺ cells with follicular centres in normal or neoplastic lymphoid tissue raises the possibility of a regulatory role of HNK-1⁺ Leu 1⁺ cells on intra-follicular B-cell proliferation.

HNK-1 antigen is expressed by large granular lymphocytes (LGL) which demonstrate spontaneous cytotoxicity against K562 cells⁹. The HNK-1⁺ fraction of peripheral blood mononuclear cells may also express a myeloid antigen M1 and subpopulations express T-cell antigens (T1, T3, T4 and T8) in varying proportions as well as HLA-DR in about 25% of the cells¹⁰. Thus the exact lineage of the NK cells has not been fully defined. The highest NK activity is associated with the HNK-1⁺ T3-M1⁺ fraction of LGL¹⁰. The HNK-1 antigen is detectable on the cell surface of LGL as well as in the cytoplasm of some LGL⁸. We were interested in determining the distribution of HNK-1⁺ cells within T- or B-cell compartments in peripheral lymphoid tissues of man. Using a biotin-avidin peroxidase system, we were able to demonstrate HNK-1⁺ cells in tissues. Figure 1a, b shows Leu 1-biotin and Leu 7-biotin treated sections of tonsillar tissue. The HNK-1⁺ cells were found clustered within the follicular centres, particularly concentrated in the region between the mantle-zone and the germinal centres. Only rarely were HNK-1⁺ cells seen in the paracortical (T-cell zone) areas or medullary cords. Identical patterns were seen in the spleen (Fig. 2a, b) and in reactive lymph nodes. Scattered HNK-1⁺ cells were seen in splenic red pulp. Few HNK-1⁺ positive cells were seen in thymic tissue. Table 1 shows the results of double-marker studies to determine the frequency of co-expression of other antigens on HNK-1⁺ cells in tonsils, lymph nodes and spleen. To avoid errors due to cross-reactions, HNK-1⁺ cells were demonstrated by Leu 7 antibody and an affinity purified, goat anti-mouse IgM (μ -chain specific), F(ab')₂ fluorescein isothiocyanate (FITC) second antibody while Leu 1, Leu 2, Leu 3 and HLA-DR antibodies were obtained in biotinylated form and used with avidin-rhodamine. The goat anti-mouse IgM, F(ab')₂ FITC did not show any fluorescence in sections treated with Leu 1, Leu 2 or Leu 3 antibodies, indicating no detectable reactivity with mouse IgG. Tonsillar, splenic and lymph node intrafollicular HNK-1⁺ cells expressed a pan-T antigen Leu 1 (equivalent to OKT1), but did not express suppressor-cytotoxic cell-associated antigen Leu 2, helper-inducer cell associated antigen Leu 3, Ia antigen (HLA-DR) or myeloid-monocytic antigen M1. The few HNK-1⁺ cells seen in T-zone

Table 1 Frequency of co-expression of other antigens on HNK-1⁺ cells

	Leu 1	% HNK-1 ⁺ cells expressing			
		Leu 2	Leu 3	HLA DR	M1
Tonsil (1)*					
Follicular	70	5	0	0	0
Paracortical	96	0	0	0	0
Lymph nodes (3)					
Follicular	97 (92-100)†	2 (0-6)	0	0	0
Paracortical	99 (96-100)	0	0	0	0
Spleen (2)					
Follicular	94 (90-98)	0	0	0	0
Red pulp	6 (4-8)	3 (2-4)	0	0	30 (25-35)

Cryostat sections of frozen tonsil, lymph nodes and spleens were incubated with 1:10 dilutions of Leu 7 antibody and biotinylated Leu 1, Leu 2, Leu 3 or HLA-DR antibodies (Becton-Dickinson) for 30 min, washed and then incubated with affinity purified goat anti-mouse IgM, F(ab')₂, FITC (Tago), 1:10 and avidin-rhodamine (Becton-Dickinson) at 1:300 for 30 min. After washing, the slides were coverslipped with a drop of glycerol-phosphate-buffered saline. Green and red fluorescence were viewed in sequence with a Zeiss fluorescence microscope fitted with sliding blue and green excitation filter sets and barrier filters. OKM1 stained preparations were incubated with affinity purified goat anti-mouse IgG (γ -chain specific), FITC (Cappel). In these preparations HNK-1⁺ cells were stained with Leu 7-biotin and avidin-rhodamine.

* Numbers in parentheses indicate the numbers of specimens studied.

† Numbers indicate means and range of values obtained.

Table 2 HNK-1⁺ cells in malignant lymphoma

Tissue	Diagnosis*	Phenotype†	% of HNK-1 ⁺ cells‡
Lymph node	Nodular, PDLL	B cell (IgM- κ)	Concentrated within nodules, 10-40
Lymph node	Nodular, histiocytic	B cell (IgG- λ)	Concentrated within nodules, 10-40
Lymph node	Diffuse, PDLL	B cell (IgM- κ)	Diffuse <10
Lymph node	Diffuse, histiocytic	B cell (IgG- κ)	Diffuse 10-20
Lymph node	Diffuse, mixed histiocytic-lymphocytic	T cell (OKT3 ⁺ , OKT4 ⁺ , OKT8 ⁻)	Diffuse 20-40
Lymph node	Diffuse, histiocytic	T cell (86% E rosette ⁺)	Rare
Tonsil	Lymphoblastic lymphoma	T cell (OKT1 ⁺ , OKT3 ⁻ , OKT6 ⁺ , OKT4 ⁺ , OKT8 ⁺)	Rare
Lymph node	Hodgkin's disease, nodular sclerosis	Polyclonal pattern	Nodules 20-40

Cell suspensions were prepared by mincing portions of lymph node biopsies in RPMI 1640+20% fetal bovine serum (FBS). Aliquots of cells were stained with FITC conjugated preparations of rabbit anti-human IgA, D, E, G, M (heavy-chain specific), κ and λ light chains¹⁶. E-rosettes were prepared as previously described¹⁶. Cryostat sections of portions of lymph node biopsies frozen in liquid nitrogen-isopentane mixtures were cut (4-6 μ m), air-dried for 2 h, fixed in 50% acetone:chloroform for 5 min at room temperature and lyophilized. Sections were rehydrated in phosphate-buffered saline containing 10% AB serum. After blocking endogenous peroxidase (3% hydrogen peroxide in water, 5 min) the sections were incubated with monoclonal antibodies OKT1, OKT3, OKT4, OKT6, OKT8 (Ortho Diagnostics) or biotinylated Leu 7 (anti-HNK-1; Becton-Dickinson) for 30 min at room temperature. After three washes, the sections were incubated with affinity purified goat anti-mouse IgG (Tago), peroxidase conjugated (for the OKT series) or avidin-peroxidase (Cedarlane) for the Leu 7 treated sections, for 30 min. After three washes, the sections were treated with diaminobenzidine and hydrogen peroxide for 5 min, washed, counterstained with haematoxylin and coverslipped.

* Diagnosis based on the Rappaport classification for non-Hodgkin's lymphomas.

† Based on cell-surface marker studies of suspensions or immunohistochemical staining of cryostat sections.

‡ Based on immunohistochemical staining of cryostat sections.

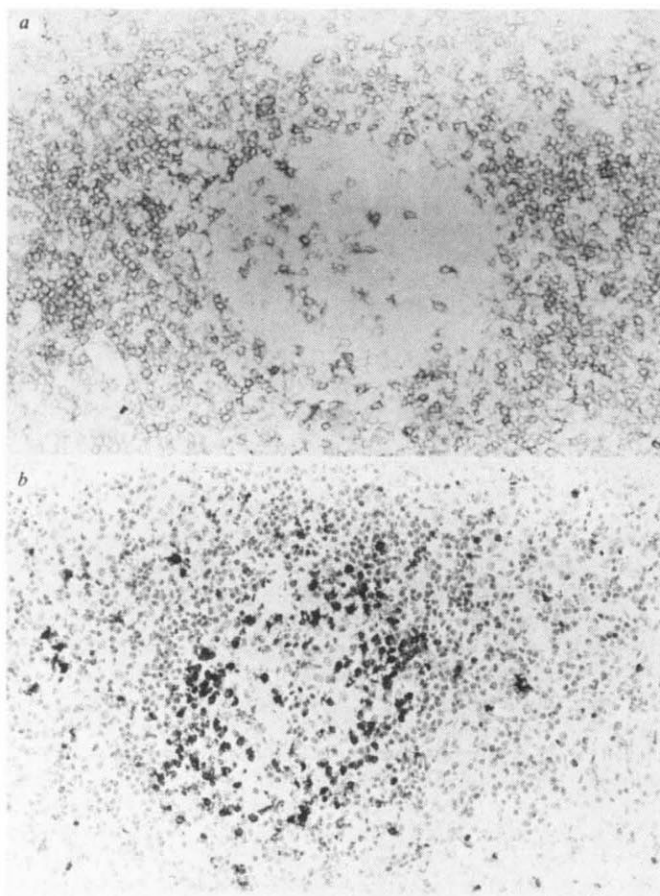


Fig. 1 *a*, Section of human tonsil treated with Leu 1-biotin and avidin-peroxidase. No counterstain. *b*, Section of human tonsil treated with Leu 7-biotin and avidin-peroxidase. Haematoxylin counterstain. $\times 75$.

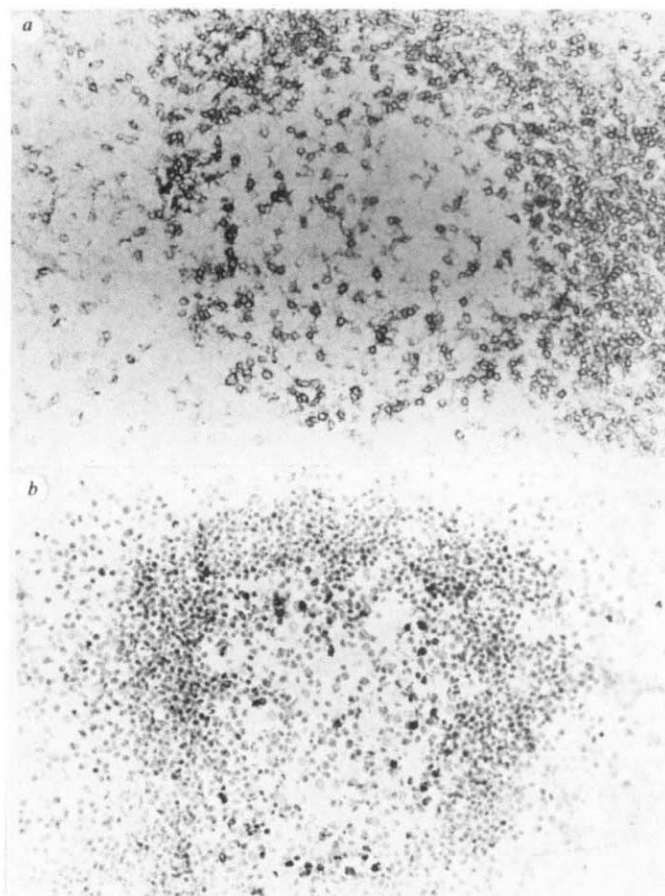


Fig. 2 *a*, Section of human spleen treated with Leu 1-biotin and avidin-peroxidase. No counterstain. $\times 50$. *b*, Section of human spleen treated with Leu 7-biotin and avidin-peroxidase. Haematoxylin counterstain. $\times 50$.

(paracortical) areas of tonsil and lymph nodes were also Leu 1⁺ Leu 2⁻ Leu 3⁻ Ia⁻ M1⁻. Splenic red pulp HNK-1⁺ cells, however, were largely Leu 1⁺ Leu 2⁻ Leu 3⁻ Ia⁻ but about 30% were M1⁺. Intrafollicular HNK-1⁺ cells were also T3⁺ (one lymph node and one tonsil tested; data not shown). Table 2 summarizes the findings in a preliminary survey of different types of malignant lymphoma. In two cases of nodular lymphoma, HNK-1⁺ cells were found clustered within the neoplastic nodules with only a few cells outside the nodules. Variable numbers of HNK-1⁺ cells were seen in diffuse lymphomas of either T or B origin, and in one case of Hodgkin's disease of nodular sclerosis type, more than 40% of the cells in the nodules were HNK-1⁺.

In vitro lysis of freshly isolated human B-cell derived leukaemia cells¹¹ or B-cell-derived lymphoma/leukaemia cell lines¹² by interferon-activated NK cells has been reported. Murine lymphomas such as the SJL/J transplantable reticulum-cell neoplasm D4, are associated with high NK activity within affected lymph nodes or spleen¹³ and a positive correlation has been reported between *in vitro* NK activity and *in vivo* resistance towards AKR lymphoma cells¹⁴. In human subjects with lymphomas of the non-Hodgkin's type, NK activity has been found to be reduced¹⁵. Our preliminary observations suggest that HNK-1⁺ cells may be closely associated with neoplastic nodules in lymphomas of follicular centre origin but may or may not be present in affected nodes in some T lymphocyte-derived lymphomas. Whether the HNK-1⁺ cells that are present in lymphomatous nodes are capable of lysing the neoplastic cells remains to be seen. Intrafollicular HNK-1⁺ cells expressed Leu 1 but not Leu 2, Leu 3 or M1, a phenotype associated with

a lower cytotoxic ability than HNK-1⁺T3⁻M1⁺ cells¹⁰. The observation that HNK-1⁺ Leu 1⁺ cells are concentrated in follicular centres suggests a possible role in intrafollicular B-cell regulation. Splenic red pulp HNK-1⁺ cells were M1⁺ and probably account for the high NK activity in spleen cell preparations. The HNK-1⁺ M1⁺ phenotype was not observed in tonsil or lymph nodes, accounting for the low NK activity in such tissues^{5,6}.

This work was supported by a grant from the National Cancer Institute of Canada. R.F.T. is the recipient of a Terry Fox Cancer Research Clerkship Award of the National Cancer Institute of Canada. We thank Sabina Pettit and Marilyn Klie for technical assistance, and Mary Jo Stone for typing the manuscript.

Received 2 December; accepted 25 May 1983.

1. Rosenberg, E. B. *et al. Int. J. Cancer* **9**, 648-658 (1972).
2. Kasai, M., Leclerc, J. C., McVay-Boudreau, L., Sheu, F. W. & Cantor, H. *J. exp. Med.* **149**, 1260-1264 (1979).
3. Loutit, J. F., Townsend, K. M. S. & Knowles, J. F. *Nature* **285**, 66 (1980).
4. Mangan, K. F., Winkelstein, A., Hartnett, M. E. & Matis, S. A. *Clin. Res.* **30**, 773A (1982).
5. Vose, B. M., Vanky, F., Argov, S. & Klein, E. *Eur. J. Immun.* **7**, 753-757 (1977).
6. Kiessling, R., Klein, E., Pross, H. & Wigzell, H. *Eur. J. Immun.* **5**, 117-121 (1975).
7. Ruco, L. P., Procopio, A., Uccini, S., Marcorelli, E. & Baroni, C. D. *Cancer Res.* **42**, 2063-2068 (1982).
8. Abo, T. & Balch, C. M. *J. Immun.* **127**, 1024-1029 (1981).
9. Abo, T., Cooper, M. D. & Balch, C. M. *Pediat. Res.* **14**, 591 (1981).
10. Abo, T., Cooper, M. D. & Balch, C. M. *J. Immun.* **129**, 1752-1757 (1982).
11. Pattengale, P. K. *et al. Int. J. Cancer* **29**, 1-7 (1982).
12. Pattengale, P. K. *et al. Int. J. Cancer* **28**, 459-468 (1981).
13. Chang, K. S. S. & Log, T. *Int. J. Cancer* **25**, 405-416 (1980).
14. Riesenfeld, I. *et al. Int. J. Cancer* **25**, 399-403 (1980).
15. Neri, A., Brugiattelli, M., Ozger, T. U. & Astaldi, G. *Boll. Ist Sieroter. Milan* **60**, 394-407 (1981).
16. Banerjee, D., Hamdy, H., Bormanis, J., Walker, T. & Richter, M. *Cancer* **41**, 1804-1810 (1978).

Three-dimensional structure, specificity and catalytic mechanism of renin

Tom Blundell, Bancinyane Lynn Sibanda
& Laurence Pearl

Laboratory of Molecular Biology, Department of Crystallography,
Birkbeck College, University of London, Malet Street,
London WC1E 7HX, UK

Renin is an aspartyl proteinase that catalyses the first, and rate-limiting, step in the conversion of angiotensinogen to the hormone angiotensin II¹. The catalysis is highly specific, and plays an important physiological part in the regulation of blood pressure². For this reason inhibitors of renin are of potential value in the treatment of certain forms of hypertension. Although progress has been made in the design of inhibitors for clinical use by modification of angiotensinogen sequences³, and as pepstatin analogues⁴ or with reduced peptide bonds⁵, we have now provided the basis for a more rational approach by the use of interactive computer graphics techniques to build a three-dimensional model of renin. The model is based on the three-dimensional structure of endothia pepsin and the primary structure of mouse renin^{6,7}, which is very similar to that of the human enzyme⁸. We show that renin may have a three-dimensional structure similar to that of other aspartyl proteinases⁹⁻¹¹.

Figure 1 shows a stereo view of the three-dimensional structure of mouse submaxillary renin constructed using interactive computer graphics. The model of endothia pepsin⁹, which has been refined by restrained least-squares at 2.1 Å resolution to an agreement factor of 16% (unpublished data), was used to place the main-chain atoms, homologous side chains, and conservatively varied side chains in closely similar orientations. This shows that almost all the main-chain atoms can assume an identical conformation in renin and that the hydrophobic core packs closely with the same volume occupation. The few insertions required, for example at the disulphide loops involving residues cystines 45, 50 and 206, 210, were easily accommodated with 'allowed' main-chain torsion angles. An extension of the β -turn in the region 279-280 accommodated the insertion of four residues including those cleaved out in the conversion of zymogen to the two-chain active enzyme^{6,7}. Finally, the more variable side chains were added in positions that optimized tertiary interactions.

The great similarity of the aspartyl proteinases so far defined by high resolution X-ray analysis, and the very few insertions and deletions in the polypeptide of renin compared with others, encourage us to believe that the model is a good representation of the three-dimensional structure. It should be of sufficient precision to recognize amino acid substitutions that are not only critical to the enzyme mechanism as far as it is defined for

other aspartyl proteinases¹²⁻¹⁴, but also affect the size and the nature—hydrophobicity, charge, and so on—of the specificity pockets. Very small and concerted changes resulting from substitutions may also be important, as they are to the understanding of the specificity of trypsin compared with kallikrein, for example¹⁵, but their definition must await a complete X-ray analysis at ~ 1.5 Å resolution.

Figure 2 shows the residues in the active site cleft, and Table 1 lists the equivalent residues in other aspartyl proteinases, which are thought to bind a substrate. These residues, first identified on the basis of model building with endothia pepsin^{16,17} and penicillopepsin¹⁸, are consistent with recent X-ray studies of enzyme inhibitor complexes^{19,20}. Figure 2 also shows the hydrogen bonding of the aspartyl proteinase residues around Asp 32 and Asp 215, defined by the high resolution refinement of endothia pepsin (unpublished data). The two aspartates have a symmetrical arrangement, and many of the conserved groups are related by a local dyad axis (unpublished data), which also relates many of the residues in the two domains^{21,22}. A similar arrangement is also found in the high resolution X-ray structure of penicillopepsin²⁵. The existence of the conserved Asp 32 and Asp 215 in renin is consistent with its inactivation by aliphatic diazo-compounds and by epoxide 1,2-epoxy-3-(*p*-nitrophenoxy)-propane in a similar manner to other aspartyl proteinases^{7,23,24}. This symmetrical environment clearly invalidates much of the earlier speculation based on the unrefined medium resolution structures concerning the identity of an aspartate with an unusually high pK ^{13,16-18}. Certain residues previously thought to have a catalytic role in all aspartyl proteinases differ in renin. Most important is Asp 304, which in renin is alanine. In other aspartyl proteinases this residue occupies a somewhat hydrophobic pocket although the carboxyl of aspartate is bound to water, which in turn binds to the surrounding water environment (unpublished data). If the mechanism of renin is in common with that of other aspartyl proteinases then this Asp 304 cannot be essential as first suggested. A similar argument applies to the otherwise invariant Asp 11 which in renin is an asparagine. We suggest that these residues might have a role mediated indirectly by hydrogen bonding in lowering the pK of the catalytically important Asp 32 and Asp 215 system in most aspartyl proteinases. In renin the optimal activity is at higher pH ⁵⁻⁸, depending on substrate²³, and so this is not necessary.

Renins are specific for angiotensinogens²⁶⁻²⁸, partial sequences of which are given in Table 2. The active sites of aspartyl proteinases bind at least seven residues, P_4 to P_3' ²⁹⁻³¹, which are accommodated in the complementary subsites indicated in Table 1 and Fig. 2. Many of the amino acids which make up specificity pockets equidistant from the scissile bond are related by the pseudo-dyad axis between the two topologically equivalent domains^{21,22}. Two such subsites at P_1 and P_1' are remarkably conserved between all aspartyl proteinases, and are mainly hydrophobic in renin as in the homologous enzymes. However, the topologically equivalent groups at 120 and 301 are valines in renin, rather than isoleucines or leucines as in the other enzymes. A preference for leucines adjacent to the

Fig. 1 A stereoview of the three-dimensional structure of mouse submaxillary renin constructed using an Evans and Sutherland Picture System II with interactive computer graphics programs by Dr T. A. Jones³⁵ and by Dr I. J. Tickle and based on the structure of the homologous aspartyl proteinase from *E. parasitica* (unpublished data). The numbering of residues follows the sequence of pepsin¹².

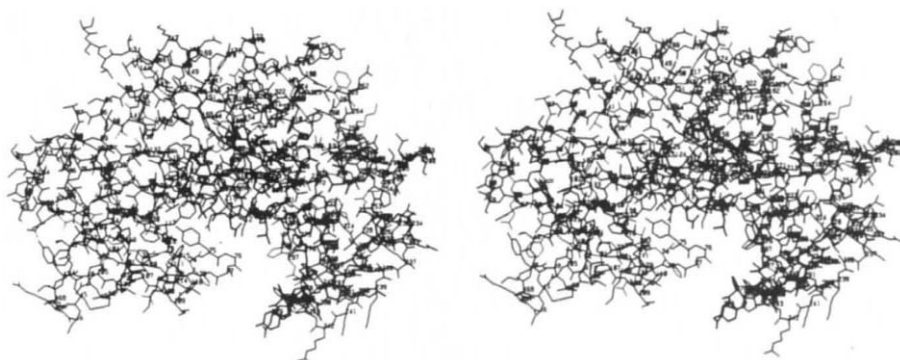


Table 1 Active site residues of aspartyl proteinases

Endothia pepsin	D	D	A	D	D	T	G	S	S	D	W	I	Y	G	D	F	T	S	T	I	D	L
Chymosin	D	S	Q	L	D	T	G	S	S	D	W	I	Y	G	T	F	T	A	E	F	D	I
Penicillopepsin	D	E	E	N	D	T	G	S	A	D	W	I	Y	G	D	F	Q	T	N	N	D	L
Porcine pepsin	D	T	E	I	D	T	G	S	S	N	W	I	Y	G	T	F	L	A	P	F	D	I
Renin	N	S	Q	I	D	T	G	S	A	N	W	I	Y	G	S	F	—	A	Q	F	D	V
Residue no. (Pepsin scheme)	11	12	13	30	32	33	34	35	36	37	39	73	75	76	77	111	112	115	116	117	118	120
Subsites		P ₃	P ₃	P ₁		P ₁		P ₂ '	P ₃ '		P ₁		P ₁		P ₂	P ₁						P ₁
					Active site aspartates																	
Residue no. (Pepsin scheme)	187	188	189	213	215	216	217	218	219	220								296	297	298	299	P ₁ ' 301
Renin	T	D	S	V	D	T	G	S	S	F	S							T	G	P	V	V
Porcine pepsin	E	G	Y	I	D	T	G	T	S	L	T							S	G	E	L	I
Penicillopepsin	Q	G	F	I	D	T	G	T	T	L	L							G	I	G	F	I
Chymosin	Q	Q	Y	I	D	T	G	T	S	K	V							S	—	Q	K	I
Endothia pepsin	Q	G	F	I	D	T	G	T	T	L	Y							G	I	G	I	I

This shows the residues in the active site clefts of endothia pepsin (V. Pedersen & L. H. Pearl, unpublished results), bovine chymosin¹⁴, porcine pepsin¹², penicillopepsin¹⁰ and mouse submaxillary renin^{6,7}. The numbering is based on the pepsin sequence. The residues are arranged so that topologically equivalent residues on the two lobes related by a pseudo-dyad axis are in the same column. Subsites are indicated according to the scheme of Schechter and Berger³¹. Key: A, Ala; D, Asp; E, Glu; F, Phe; G, Gly; I, Ile; K, Lys; L, Leu; N, Asn; P, Pro; Q, Gln; S, Ser; T, Thr; V, Val; W, Trp; Y, Tyr.

scissile bond in substrates does not seem to be reflected at subsites P₁ and P₁' in renin and this is consistent with the observations of Burton *et al.*³ that inhibitors with phenylalanines at S₁ and S₁' bind well.

The residues in subsite P₃ are also similar; however, residue 189 in subsite P₃' is remarkably different. In all other aspartyl proteinase sequences this residue is large and aromatic, either phenylalanine or tyrosine, whereas in renin it is a serine, which is small and hydrophilic. This may account for the fact that renin prefers large groups at S₃', either tyrosine or histidine, which may hydrogen bond to the serine.

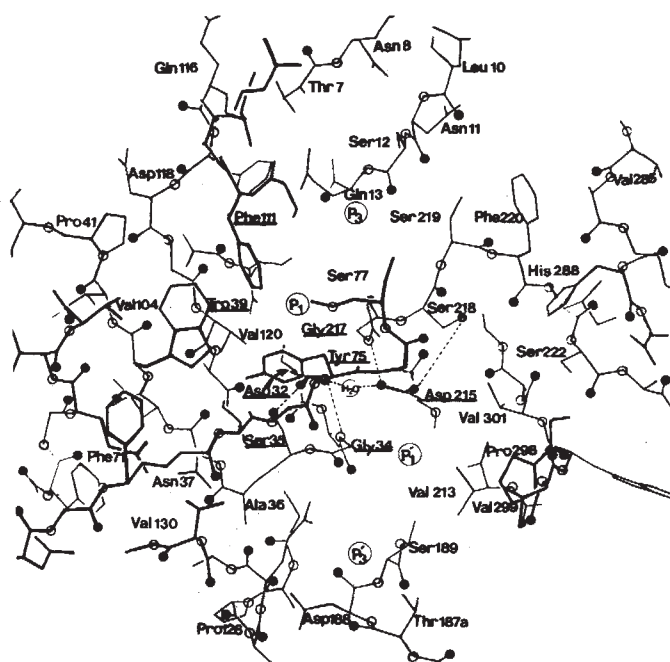


Fig. 2 The residues in the renin three-dimensional structure which form the inner surface of the active site cleft viewed from the top. The specificity subsites P₃, P₁, P₁' and P₃' (after ref. 31) are defined in Table 2. Residues which are invariant amongst aspartyl proteinases are underlined.

At the edges of the active site cleft there are some remarkable differences between renin and the homologous enzymes. These are illustrated in Fig. 3. They include a surface loop with proline residues at 293, 294, 295 and 298. The external surface of the 'flap', residues 71–83, which lies across the cleft, becomes highly basic with Lys 81, Arg 79 and His 74. There is a further surface region involving Lys 239, Lys 241, Arg 242 and His 244; none of these residues is basic in any other aspartyl proteinase. In addition, in the zymogen there is a β -loop easily available on the surface containing three basic groups from which two arginines are cleaved in the conversion to an active enzyme.

This unusual surface region may have an important functional role. First it may facilitate recognition of the renin zymogen by the important conversion enzymes. The resulting residues and the other surface loops may then have a role in binding angiotensinogen, which in rat is a protein of molecular weight

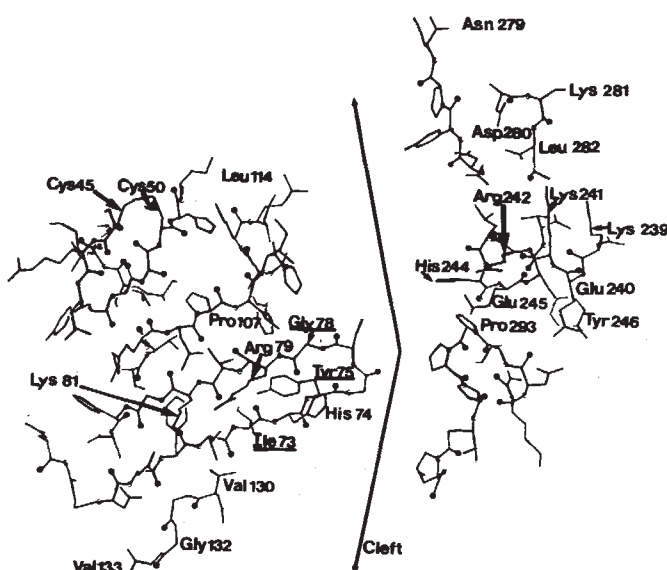


Fig. 3 The arrangement of the residues on the edge of the active site cleft of renin, viewed in the same direction as Fig. 2. Residues which are invariant amongst aspartyl proteinases are underlined.

Table 2 Active site-binding sequences of angiotensinogens

Enzyme	P ₄	P ₃	P ₂	P ₁	P ₁ '	P ₂ '	P ₃ '
Substrate	S ₄	S ₃	S ₂	S ₁	S ₁ '	S ₂ '	S ₃ '
Horse	Pro	Phe	His	Leu	Leu	Val	Tyr
Human	Pro	Phe	His	Leu	Val	Ile	His
Rat	Pro	Phe	His	Leu	Leu	Tyr	Tyr

The sequences of horse²⁶, human²⁷ and rat²⁸ angiotensinogens which bind to the active site cleft of renin. The residues (S_n) and the complementary enzyme subsites (P_n) are labelled according to the scheme of Schechter and Berger³¹.

~57,000 (ref. 28); it would be surprising if renin were specific only for residues in the locality of the scissile bond.

We are now using computer graphics programs³²⁻³⁴ to investigate the interaction of putative substrates and inhibitors with a view to designing molecules which might be more effective in the treatment of hypertension. Of particular value in our model is the inferred conformation of the bound substrate and the probable size and chemical nature of the specificity pockets. When the sequence of the human renin becomes available, probably as a cDNA sequence, the same approach can be used to further define the small differences of specificity which will enhance the predictive value of this modelling technique.

We thank Professor T. Inagami for sending us a preprint of his paper on the amino acid sequence of renin; Dr Ian Tickle for the development of computer graphics software; the UK SERC for the support of the Evans and Sutherland Picture System II; and Professors D. M. Blow and A. C. T. North for their helpful comments.

Received 10 February; accepted 2 June 1983.

- Skeggs, L. T. *et al.* in *The Renin-Angiotensin System* (eds Johnson, J. A. & Anderson, R. R.) 1-27 (Plenum, New York, 1980).
- Atkinson, A. B. *et al.* in *The Therapeutics of Hypertension*, 29-61 (Academic and The Royal Society of Medicine, London, 1980).
- Burton, J., Cody, R. J., Herd, J. A. & Haber, E. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5476-5479 (1980).
- Boger, J. *et al.* *Nature* **303**, 81-84 (1983).
- Szelke, M. *et al.* *Nature* **299**, 555-557 (1982).
- Panthier, J.-J., Foote, S., Chambrault, B., Strosberg, P. V. & Rougeon, F. *Nature* **298**, 90-92 (1982).
- Misono, K. S., Chang, J.-J. & Inagami, T. *Proc. natn. Acad. Sci. U.S.A.* **79**, 4858-4862 (1982).
- Hirose, S., Workman, R. J. & Inagami, T. *Circulation Res.* **45**, 275-279 (1977).
- Subramanian, E. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **74**, 559-566 (1977).
- Hsu, I. N. *et al.* *Nature* **266**, 140-145 (1977).
- Andreeva, N. S. *et al.* *J. molec. Biol. (Russian)* **12**, 922-935 (1978).
- Tang, J. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **70**, 3437-3439 (1973).
- James, M. N. G., Delbaere, L. T. J. & Hsu, I. N. *Nature* **267**, 808-813 (1977).
- Foltmann, B. & Pedersen, V. B. *Adv. exp. med. Biol.* **95**, 3-22 (1977).
- Bode, W. *et al.* *J. molec. Biol.* **164**, 237-282 (1983).
- Blundell, T. L. *et al.* *FEBS Lett.* **52**, 81-94 (1978).
- Blundell, T. L. *et al.* in *Enzyme Regulation and Mechanism of Action* (eds Mildner, P. & Ries, B.) 281-288 (Pergamon, Oxford, 1980).
- James, M. N. G., Hsu, I. N., Hofmann, T. & Sielecki, A. in *Structural Studies on Molecules of Biological Interest* (eds Dodson, G., Glusker, J. P. & Sayre, D.) 350-389 (Clarendon, Oxford, 1981).
- James, M. N. G., Sielecki, A., Salituro, F., Rich, D. H. & Hofmann, T. *Proc. natn. Acad. Sci. U.S.A.* **79**, 6137-6141 (1982).
- Bott, R., Subramanian, E. & Davies, D. R. *Biochemistry* **21**, 6956-6962 (1982).
- Tang, J., James, M. N. G., Hsu, I. N., Jenkins, J. A. & Blundell, T. L. *Nature* **271**, 618-621 (1978).
- Blundell, T. L., Sewell, B. T., & McLachlan, A. D. *Biochim. biophys. Acta* **580**, 24-31 (1979).
- Inagami, T., Murakami, K., Misono, K., Workman, R. J., Cohen, S. & Suketa, Y. in *Acid Proteinases, Structure, Function and Biology* (ed. Tang, J.) 225-247 (Plenum, New York, 1977).
- Misono, K. S. & Inagami, T. *Biochemistry* **19**, 2616-2622 (1980).
- James, M. N. G. & Sielecki, A. R. *J. molec. Biol.* **163**, 299-361 (1983).
- Skeggs, L. T., Khan, J. R., Lentz, K. & Shumway, N. P. *J. exp. Med.* **106**, 439-453 (1957).
- Tewksbury, D. A., Dart, R. A. & Travis, J. *Biochem. biophys. Res. Commun.* **99**, 1311-1315 (1981).
- Bouhnik, J. *et al.* *Biochemistry* **20**, 7010-7015 (1981).
- Fruton, J. S. *Adv. Enzym.* **44**, 1-36 (1976).
- Powers, J. C., Harley, A. D. & Myers, D. V. in *Acid Proteinases, Structure, Function and Biology* (ed. Tang, J.) 141-157 (Plenum, New York, 1977).
- Schechter, I. & Berger, A., *Biochem. biophys. Res. Commun.* **27**, 157-162 (1967).
- Busetta, B., Tickle, I. J. & Blundell, T. L. *J. appl. Crystallogr.* (in the press).
- Honegger, A. & Blundell, T. L. in *Insulin-like Growth Factors* (ed. Spencer, M.) (de Gruyter, Berlin, in the press).
- Pearl, L. H. & Honegger, A. *J. molec. Graphics* **1**, 9-12 (1983).
- Jones, T. A. *J. appl. Crystallogr.* **11**, 268-272 (1978).

Retrovirus genomes methylated by mammalian but not bacterial methylase are non-infectious

D. Simon*, H. Stuhlmann†, D. Jähner†, H. Wagner*, E. Werner* & R. Jaenisch†

* Robert Koch-Institut, Abteilung Biochemie, Nordufer 20, D-1000 Berlin (West) 65, FRG

† Heinrich-Pette-Institut für Experimentelle Virologie und Immunologie an der Universität Hamburg, Martinistrasse 52, D-2000 Hamburg 20, FRG

The biological importance of DNA methylation for gene expression in eukaryotes is becoming increasingly evident^{1,2}, and a direct role of methylation in gene expression has been suggested by an analysis of the infectivity of integrated retroviral genomes in a transfection assay³⁻⁵. These studies, however, did not address whether specific methylatable residues are involved in gene regulation. Methylation by sequence-specific bacterial DNA methylases has been shown to suppress the expression of some genes⁶⁻⁹, but not others¹⁰. To investigate the effect of methylation on gene expression without having to rely on sequence-specific methylases a rat liver enzyme was used to methylate *in vitro* all C-G dinucleotides of a proviral genomic clone. This treatment reduced the biological activity of Moloney murine leukaemia virus (M-MuLV) proviral DNA by more than three orders of magnitude, whereas complete methylation of 35 *HpaII* sites in the same DNA had only a marginal effect. The rat methylase-induced inactivation was reversible, as treatment of recipient cells with 5-azacytidine rendered the non-infectious viral genomes biologically active. This suggests that methylation in other C-G dinucleotides than those detectable with restriction enzymes can be crucial for gene expression.

The proviral genomes were derived from Mov-3 and Mov-9 mice, which carry M-MuLV proviral DNA in their germ line and regularly express infectious virus^{11,12}. The clones pMov-3 and pMov-9 were derived as *EcoRI* fragments consisting of the viral genome with 5-8-kilobase (kb) flanking mouse sequences (Fig. 1A)^{4,5}. In previous studies we have shown that these clones are highly infectious when transfected onto 3T3 cells^{4,5}. The methylase used in this study was isolated from regenerating rat liver¹³ and purified 3,000 to 5,000-fold (Fig. 1). Characterization of enzymatic activity revealed that this methylase accepts as substrates single-stranded as well as double-stranded DNA. The dinucleotide C-G is the preferred recognition sequence for *in vitro* methylation, but methyl groups are also incorporated into a fraction of C-A and C-T dinucleotides¹⁴.

Chemical analysis of *in vitro* methylated DNA revealed that more than 95% of all C-G and less than 15% of the C-T and C-A dinucleotides present in pMov-3 were methylated (Table 1). Restriction enzyme analysis showed that the methylated DNA was almost completely resistant to digestion with methylation-sensitive enzymes *HpaII* and *HhaI* (Fig. 1B). Together, these enzymes have more than 50 recognition sites in pMov-3, and more than 50 in pBR322^{15,16}. These results confirm that the rat liver enzyme *do novo* methylates efficiently the C-G dinucleotides present in the retroviral clones.

Unmethylated and *in vitro* methylated DNA was tested for biological activity. The results in Table 2 show that 5 ng of mock-methylated pMov-3 or pMov-9 DNA induced approximately 110-140 XC plaques when transfected onto 3T3 cells, in agreement with previous results^{4,5}. When the same DNA was incubated with the rat liver methylase, no XC plaques were induced by 40 ng of DNA. This result shows that methylation reduced the biological activity of the retroviral clones by more than three orders of magnitude.

We have shown previously that *in vivo* methylated M-MuLV proviral genomes are non-infectious^{17,18}. However, when recipient 3T3 cells are treated with 5-azacytidine, which is

thought to inhibit maintenance methylation¹⁹, infectivity of the retroviral DNA is restored¹⁷. To test whether the *in vitro* methylated pMov-9 DNA was potentially infectious, 3T3 cells were transfected with methylated pMov-9 DNA and 24 h later exposed to 30 μ M 5-azacytidine for 1 day. Table 3 shows that infectivity was observed in three different experiments, indicating that *in vitro* methylation did not irreversibly inactivate the DNA. The transfected DNA probably integrates into the cel-

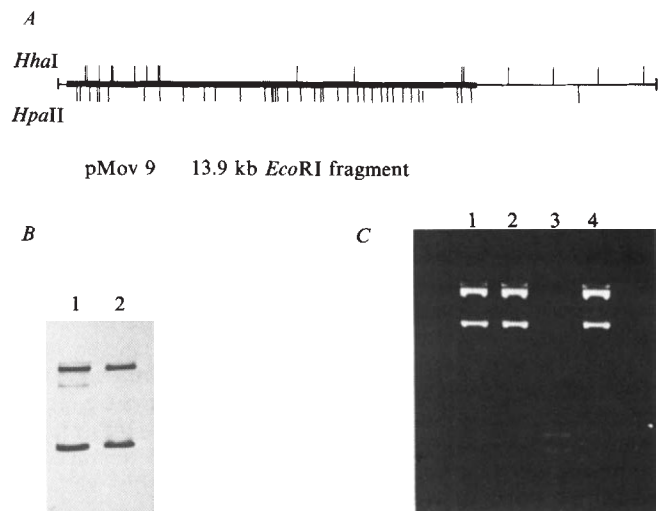


Fig. 1 **A**, *HhaI* and *HpaII* restriction map of the 13.9-kb pMov-9 *EcoRI* fragment. The black bar represents the M-MuLV provirus (8.8 kb) flanked by a 0.4-kb mouse DNA at the 5' end and 4.0 kb at the 3' end⁵. **B**, Digestion of pMov-9 DNA, which was methylated by the liver enzyme, with *HpaII* and *HhaI*. Lane 1, methylated pMov-9 DNA digested first with *HpaII* and *HhaI*, and then with *EcoRI*; lane 2, control DNA digested only with *EcoRI*. The two bands correspond to the *EcoRI* fragments of Mov-9 (13.9 kb) and pBR322 (4.4 kb). **C**, Digestion of *HpaII*-methylated pMov-9 DNA with *HpaII* endonuclease. Lane 1, methylated DNA, *HpaII*-digested; lane 2, methylated DNA, undigested; lane 3, mock-methylated DNA, *HpaII*-digested; and lane 4, mock-methylated DNA, undigested. The two bands in lanes 1, 2 and 4 correspond to forms I and II of pMov-9.

Methods: **B**, pMov-9 DNA (800 ng) was methylated for 60 h at 30 °C in a total volume of 800 μ l containing 20 mM Tris, pH 7.8, 60 mM KCl, 5 mM EDTA, 0.5 mM dithiothreitol (DTT), 10% glycerol, 0.05% Triton X-100, 200 μ g ml⁻¹ bovine serum albumin (BSA), 5 μ M ³H-SAM (15 Ci mmol⁻¹) and 800 units of DNA methylase from rat liver. The enzyme was purified from 120 g regenerating rat liver as described previously¹³ with the following modifications. The enzyme-containing fractions from the DEAE-Sephadex A50 gradient, after twofold dilution with buffer A (20 mM Tris, pH 7.8, 1 mM EDTA, 0.5 mM DTT), were stirred with heparin-Sepharose (Pharmacia) for 1 h and the enzyme eluted in a column (0.9 cm diameter) with a 60 ml linear gradient of from 85 to 320 mM KCl in buffer A plus 10% glycerol. After precipitation with ammonium sulphate (40–60%) 4,000 units (~1 mg protein) were chromatographed on a Sephadex G-150 column (0.9 $\times$ 60 cm) in buffer A plus 0.6 M KCl and 15% glycerol. The active fractions were finally concentrated 10-fold on an Amicon UM 20 membrane. After the methylation reaction the mixture was deproteinized. The extent of methylation was monitored by digestion with a 50-fold excess of *HpaII* and *HhaI*. After digestion with *EcoRI* the DNA was analysed on a 0.7% agarose gel (lane 1). The same amount of control DNA (lane 2) was digested with *EcoRI* only. Electrophoresis was followed by fluorography with ³H-enhancer (NEN). **C**, pMov-9 DNA (700 ng) was methylated for 16 h at 33 °C in a total volume of 100 μ l containing 50 mM Tris, pH 7.8, 5 mM EDTA, 1 mM DTT, 5% glycerol, 200 μ g ml⁻¹ BSA, 30 μ M and 0.5 units *HpaII* methylase (1 unit incorporates 10 pmol CH₃ in 15 min at 37 °C). As control, DNA was mock-methylated in the absence of SAM. The *HpaII* methylase was purified in six steps as published elsewhere^{32,33}. The methylated DNA was deproteinized and digested with a 20-fold excess of *HpaII* endonuclease. 200 ng DNA per lane were separated on a 0.7% agarose gel in the presence of ethidium bromide.

lular genome directly on transfection and does not replicate autonomously in the cell²⁰. We therefore cannot completely exclude that the low level of non-C-G methylation introduced by the rat liver methylase has some effect on the biological activity of the retroviral genomes²¹. However, since the maintenance methylase has been shown specifically to methylate C-G dinucleotides²², the restoration of biological activity by 5-azacytidine treatment strongly suggests that inactivation of the M-MuLV genomes was due to methylation of C-G dinucleotides.

When the retroviral clones were methylated with *HpaII* methylase, all *HpaII* sites were resistant to cleavage by *HpaII* (Fig. 1C). Transfection of this DNA onto 3T3 cells resulted in a reduction of infectivity by only 40–70%. The more than 1,000-fold reduction of infectivity by methylation with the rat liver enzyme (Table 2) contrasts with this result, and indicates that DNA methylation is causally involved in gene regulation.

In agreement with our results, Hoffmann *et al.*¹⁰ have shown that *in vitro* methylation of M-MuLV genomes at *HpaII* and *HhaI* sites had no effect on biological activity. *In vivo*, however, a strong correlation between virus gene expression, proviral DNA infectivity and methylation was observed in animals carrying endogenous M-MuLV genomes^{3,4}. Furthermore, we have shown that *de novo* methylation of retroviral genomes introduced into mouse embryonal cells completely abolishes biological activity^{17,18}. In all instances when the viral genomes were methylated *in vivo* the inactivation of biological activity strictly correlated with methylation at *HpaII* and *HhaI* sites. Thus the effect of *in vivo* methylation on biological activity was mimicked *in vitro* by the rat liver methylase, but not by the bacterial methylases. Our results therefore strongly suggest that *in vivo* methylation modifies cytosine residues in more sites than detectable by standard restriction enzyme analysis²³. They indicate that sequences other than *HpaII* and *HhaI* sites are crucial for control of retroviral genome expression. With other genes, however, *in vitro* methylation by *HpaII* methylase was shown to have a direct effect on biological activity^{6–9}. This suggests that these sites can be important for regulation of some genes, and recent evidence has been obtained that such sequences may be located at the 5' end of genes^{24–27}.

The molecular mechanisms of how DNA methylation may affect gene expression have not yet been established. Methylation of C-G dinucleotides has been shown to alter the interactions of regulatory proteins with DNA²⁸, and to cause a polymer

Table 1 Distribution of methylation among 5' cytosine dinucleotides in methylated pMov-3 DNA

	C-G	C-A	C-T	C-C
Methylation of dinucleotides (d.p.m.)	18,740	5,550	2,150	0
Fraction of dinucleotides methylated (%)	98.6	13.8	6.0	0

75 ng Mov-3 DNA were methylated with 75 units of purified rat liver DNA methylase and ³H-S-adenosyl-L-methionine (SAM). The dinucleotide analysis was performed essentially as described elsewhere^{14,23}. The methylated DNA was digested with 30 μ g DNase I after addition of 150 μ g carrier DNA in 10 mM Tris, pH 7.5, 8 mM MgCl₂, 2 mM CaCl₂ for 16 h at 37 °C. After dephosphorylation with alkaline phosphatase (calf intestine, Boehringer) total dinucleotides were purified from oligonucleotides on a DEAE-Sephadex column (10 ml), concentrated on 2 mg activated charcoal and separated by high voltage paper electrophoresis in 5% pyridine-acetate buffer, pH 3.4. The spots of ³H-methylcytosine containing dinucleotides were eluted and the isomers (for example mC-G and G-mC) distinguished by treatment with snake venom phosphodiesterase (Worthington) and separation of the products by high voltage paper electrophoresis in 50 mM Na citrate buffer, pH 2.5. The overall methylation was 4.2%, of which 71% was found in C-G dinucleotides. The fraction of C-G dinucleotides present in pMov-3 DNA is approximately 3.1%, as calculated from published sequences (3.1% C-G in 9 kb M-MuLV DNA¹⁵, 7% in 4.4 kb pBR322 DNA¹⁶ and about 0.9% in flanking mouse DNA³¹). Methylation of C-A and C-T was calculated in the same way.

Table 2 Infectivity of cloned retroviral genomes after *in vitro* methylation by rat liver methylase

DNA	<i>In vitro</i> methylation	Aza-C treatment of recipient cells	Total no. of XC plaques
pMov-3 10 ng	+	—	0
5 ng	—	—	116
pMov-9 40 ng	+	—	0
40 ng	+	+	101
5 ng	—	—	137
10 ng	—	—	438
40 ng	—	—	550
5 ng	—	+	165

Cloned plasmid DNA (pMov-3 and pMov-9) was *in vitro* methylated with rat liver methylase as described in Table 1 legend. *In vitro* methylated DNA was linearized by cleavage with *EcoRI* and tested for infectivity by transfection of NIH 3T3 cells as described elsewhere³. The extent of *in vitro* methylation was assayed by cleavage of an aliquot with *HhaI* and *HpaII* and the sequence specificity by dinucleotide analysis (see Fig. 1). A DNA sample used for transfection contained 5–40 ng plasmid DNA and 30 µg of BALB/c carrier DNA. As a control, DNA mock-methylated by incubation without SAM was tested for infectivity. In some experiments NIH 3T3 receptor cells were treated for 24 h with 30 µM 5-azacytidine (Aza-C) 24 h after addition of transfected DNA. The recipient NIH 3T3 cells were tested for virus production by the XC plaque assay after reaching confluency. The number of XC plaques is given as the mean of two or three independent transfection assays.

Table 3 Infectivity of retroviral genomes after *in vitro* methylation by *HpaII* methylase

DNA	<i>In vitro</i> methylation	Aza-C treatment of recipient cells	XC plaques
pMov-9 5 ng	+	—	24
50 ng	+	—	289
20 ng	+	+	45
5 ng	—	—	80
50 ng	—	—	450

pMov-9 DNA was *in vitro* methylated by *HpaII* methylase (Fig. 1C). An aliquot of *HpaII*-methylated DNA was cut with *HpaII* to determine the extent of methylation of CCGG. *In vitro* methylated DNA was linearized with *EcoRI* and tested for infectivity by DNA transfection.

to adopt the left-handed or Z-DNA conformation in physiological salt conditions²⁹. Our results clearly indicate that restriction enzyme analysis is insufficient to identify sites of methylation that are important for gene regulation. The rat liver methylase, however, permits a functional analysis independent of sequence specificity. The availability of marker rescue assays to map mutations in retroviral genomes³⁰ should enable us to define crucial target regions for methylation in proviral genomes.

This work was supported by grants from the Deutsche Forschungsgemeinschaft and the Stiftung Volkswagenwerk. The Heinrich-Pette-Institut is supported by Freie und Hansestadt Hamburg and by Bundesministerium für Jugend, Familie und Gesundheit.

Received 30 March; accepted 6 June 1983.

- Ehrlich, M. & Wang, R. *Science* **212**, 1350–1357 (1981).
- Felsenfeld, G. & McGhee, J. *Nature* **296**, 602–603 (1982).
- Stuhlmann, H., Jähner, D. & Jaenisch, R. *Cell* **26**, 221–232 (1981).
- Harbers, K., Schnieke, A., Stuhlmann, H., Jähner, D. & Jaenisch, R. *Proc. natn. Acad. Sci. U.S.A.* **78**, 7609–7613 (1981).
- Chumakov, I., Stuhlmann, H., Harbers, K. & Jaenisch, R. *J. Virol.* **42**, 1088–1098 (1982).
- Wigler, M., Levy, D. & Perucho, M. *Cell* **24**, 33–40 (1981).
- Stein, R., Razin, A. & Cedar, H. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3418–3422 (1982).
- Vardimon, L., Kressmann, A., Cedar, H., Maechler, M. & Doerfler, W. *Proc. natn. Acad. Sci. U.S.A.* **79**, 1073–1077 (1982).
- McGeady, M. L., Jhappan, C., Ascione, R. & Vande Woude, G. F. *Molec. cell. Biol.* **3**, 305–314 (1983).
- Hoffmann, J. et al. *J. Virol.* **44**, 144–157 (1982).
- Jähner, D. & Jaenisch, R. *Nature* **287**, 456–458 (1980).

- Jaenisch, R. et al. *Cell* **24**, 519–529 (1981).
- Simon, D., Grunert, F., Acken, K., Döring, H. & Kroger, H. *Nucleic Acids Res.* **5**, 2153–2167 (1978).
- Simon, D. et al. in *Biochemistry of S-Adenosylmethionine and Related Compounds* (eds Borchardt, T. & Creveling, C.) 267–273 (Macmillan, London, 1982).
- Shinnick, T., Lerner, R. & Sutcliffe, J. *Nature* **293**, 543–548 (1981).
- Sutcliffe, J. G. *Cold Spring Harb. Symp. quant. Biol.* **43**, 77–90 (1979).
- Stewart, C., Stuhlmann, H., Jähner, D. & Jaenisch, R. *Proc. natn. Acad. Sci. U.S.A.* **79**, 4098–4102 (1982).
- Jähner, D. et al. *Nature* **298**, 623–628 (1982).
- Creusot, F., Acs, G. & Christman, J. *J. biol. Chem.* **257**, 2041–2048 (1982).
- Copeland, N., Zeienetz, A. & Cooper, G. *Cell* **17**, 993–1002 (1979).
- Waechler, D. & Baserga, R. *Proc. natn. Acad. Sci. U.S.A.* **79**, 1106–1110 (1982).
- Gruenbaum, J., Cedar, H. & Razin, A. *Nature* **295**, 620–622 (1982).
- Sano, H. & Sager, R. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3584–3588 (1982).
- Bird, A., Taggart, M. & Macleod, D. *Cell* **26**, 381–390 (1981).
- Sweet, R., Chao, M. & Axel, R. *Cell* **31**, 347–353 (1982).
- Ott, M. et al. *Cell* **30**, 825–833 (1982).
- Kruczek, I. & Doerfler, W. *EMBO J.* **1**, 409–414 (1982).
- Fisher, E. & Caruthers, M. *Nucleic Acids Res.* **7**, 401–408 (1979).
- Behe, M. & Felsenfeld, G. *Proc. natn. Acad. Sci. U.S.A.* **78**, 1619–1623 (1981).
- Schnieke, A., Stuhlmann, H., Harbers, K., Chumakov, I. & Jaenisch, R. *J. Virol.* **45**, 503–513 (1983).
- Schwartz, M., Trautner, T. & Kornberg, A. *J. biol. Chem.* **237**, 1961–1967 (1962).
- Mann, M. B. & Smith, H. O. *Nucleic Acids Res.* **4**, 4211–4221 (1977).
- Yoo, O. J. & Agarwal, K. L. *J. biol. Chem.* **255**, 6445–6449 (1980).

Transcription of the *KpnI* families of long interspersed DNAs in human cells

Bridget Shafit-Zagardo, Fred L. Brown, Paul J. Zavodny & Joseph J. Maio

Department of Cell Biology, Albert Einstein College of Medicine, Bronx, New York 10461, USA

The mammalian genome contains a variety of interspersed repetitive sequences of unknown function. It has, however, been suggested that interspersed repetitive sequences and their RNA transcripts are involved in the coordinate regulation of gene expression¹. Two major families of interspersed sequences in primates are the so-called *Alu* and *KpnI* families^{2–13}. Members of the *KpnI* families range in length from 1.2 to over 6 kilobases (kb). They exist in generally clustered arrangements, in 6×10^4 to 10^5 copies per diploid genome. Something is known of the arrangements of *KpnI* family sequences near human structural genes^{4,7,11,12}, but there has been no information on transcription of the sequences. We report here that the *KpnI* sequences are transcribed in HeLa cells by RNA polymerase II into abundant and heterogeneous species of RNA. The transcripts range in length from about 200 bases to over 5 kb, and are found predominantly in the non-polyadenylated fraction of the nuclear RNA. Transcripts homologous to both of the complementary strands of the *KpnI* sequences are present, but with a strong bias towards one strand.

KpnI sequences were isolated from the conspicuous bands formed upon gel electrophoresis of *KpnI* digests of human DNA, and cloned into the *KpnI* site of the plasmid pBK5⁸. Analysis of 24 different cloned *KpnI* sequences indicated the existence of at least six distinct families, defined by shared restriction sites and cross-hybridization between members of a family.

In the studies reported here, plasmids representative of six distinct *KpnI* families were used: pBK(1.2)11, pBK(1.5)54, pBK(1.8)26, pBK(1.8)11, pBK(1.8)20 and pBK(1.9)51 (the numbers in parentheses refer to the length in kb of the corresponding genomic segments). The different *KpnI* family cloned inserts were radioactively labelled by nick-translation and hybridized to HeLa cell nuclear and cytoplasmic RNAs immobilized on nitrocellulose filters, as shown for the *KpnI* 1.2 and 1.5 kb probes in Fig. 1. Several features of the *KpnI* families are illustrated by Fig. 1:

(1) The *KpnI* 1.2 and 1.5 kb sequences are transcribed into nuclear RNAs (lanes a–c, g–i), and the transcripts are located predominantly in the poly(A)[–] fraction of the nuclear RNA (lanes c, i). Hybridization was observed with the poly(A)⁺ nuclear RNA fractions that bound to oligo(dT)–Sephadex

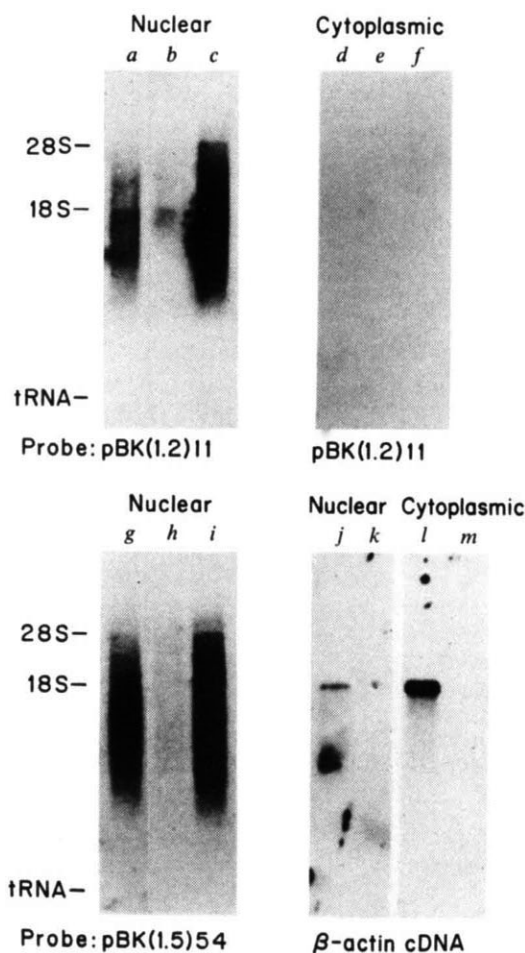


Fig. 1 Hybridization of nick-translated *KpnI* family sequences with nuclear and cytoplasmic HeLa RNA subfractions. In lanes *a-f* the nick-translated probe was the pBK(1.2)11 insert, as indicated. Lane *a*, total nuclear RNA; *b*, nuclear poly(A)⁺ RNA; *c*, nuclear poly(A)⁻ RNA; *d*, total cytoplasmic RNA; *e*, cytoplasmic poly(A)⁺ RNA; *f*, cytoplasmic poly(A)⁻ RNA. Lanes *g-i* show the nuclear total, poly(A)⁺ and poly(A)⁻ RNA fractions hybridized to the pBK(1.5)54 nick-translated insert. In lanes *j-m*, the β -actin cDNA probe was hybridized to nuclear poly(A)⁺ RNA (*j*); nuclear poly(A)⁻ RNA (*k*); cytoplasmic poly(A)⁺ RNA (*l*) and cytoplasmic poly(A)⁻ RNA (*m*).

Methods: Approximately 10^9 HeLa cells grown in suspension culture were washed twice in 200 ml of lysis buffer (0.15 M NaCl, 1.5 mM MgCl₂, 0.01 M Tris-HCl, pH 7.4) and then resuspended in 10 ml of lysis buffer containing 10 mM vanadyl-ribonucleotide complexes¹⁸ and 0.5% NP40 (Shell Chemicals). The nuclei were released by Dounce homogenization and pelleted by centrifugation at 1,000 g for 10 min at 4°C. The supernatant containing the cytoplasmic RNA was removed, and an equal volume of 2× acetate buffer added (1× buffer is 100 mM Na-acetate, pH 5.2, 1 mM EDTA). The nuclei were resuspended in lysis buffer without NP40, layered over 2 volumes of 24% (w/v) sucrose in lysis buffer, and pelleted as described above. The pellet was resuspended in 10 ml of 1× acetate buffer. SDS was then added to both the nuclear and cytoplasmic fractions to 0.5% before phenol extractions at 65°C¹⁹. The extractions were repeated until the presence of the vanadyl complexes was no longer evident. The RNA fractions were then extracted twice with chloroform/isoamyl alcohol (24:1) at room temperature and stored in ethanol at -20°C. To eliminate DNA contamination, the RNA samples were subjected to LiCl precipitation before use²⁰. The nuclear and cytoplasmic poly(A)⁺ and poly(A)⁻ fractions were separated by oligo(dT) column chromatography as previously described²¹. For electrophoresis, 10 µg of RNA were applied per lane in 1% agarose denaturing gels containing 50% formamide-6% formaldehyde (v/v). Gel electrophoresis, nitrocellulose blotting and hybridization were according to Thomas²² and Wahl *et al.*²³. All hybridization mixtures contained 10⁷ c.p.m. of probe labelled to approximately 10⁸ c.p.m. µg⁻¹ by nick-translation. The filters were hybridized for 24 h in 50% (v/v) formamide-buffer²² at 42°C, and then washed four times in 2×SSC, 0.1% SDS at room temperature, once at 55°C and once in 0.1×SSC, 0.1% SDS at 65°C (15 min per wash). The autoradiograms were exposed for 2-7 days at -70°C using sensitized Kodak X-Omat film and an intensifying screen. From microdensitometry, the intensity of the autoradiograms was reduced over 90% by treatment of the nuclear RNA with-pancreatic and T₁ ribonucleases (100 U ml⁻¹ in 0.2 M NaCl at 37°C for 15 min) before electrophoresis.

(lanes *b*, *h*) or to poly(U)-Sephacrose columns. Microdensitometric measurements showed that the autoradiographic signals obtained with the nuclear poly(A)⁻ fractions were three- to fourfold more intense than those obtained with the nuclear poly(A)⁺ fractions. If appropriate corrections are made for the total amounts of nuclear RNA present in the poly(A)⁻ and poly(A)⁺ fractions, these results then indicate that over 90% of the *KpnI* transcripts are present in the nuclear poly(A)⁻ RNA. Thus, under steady-state conditions, the bulk of the various *KpnI* family transcripts are either not polyadenylated, or their poly(A) tails, if present, are too short to bind efficiently to these affinity columns. Abundant nuclear RNA transcripts similar to those observed with the *KpnI* 1.2 and 1.5 kb probes were also detected with the probes pBK(1.8)26 and pBK(1.8)11.

However, it appears that sequences homologous to the *KpnI* family represented by the insert of plasmid pBK(1.8)20 are not transcribed into specific nuclear RNAs in HeLa cells (data not shown). Further, no transcripts of this family were detected in the *in vivo* or *in vitro* RNA transcription systems described below (Figs 2A, 3A).

(2) The transcripts are present in HeLa cells as heterogeneous populations of molecules ranging in length from about 200 bases to over 5 kb. The presence of discrete bands superimposed on the background of heterogeneous RNAs suggested that some transcripts containing *KpnI* sequences are more abundant than others in the nucleus.

(3) The *KpnI* family transcripts are largely, if not exclusively confined to the nucleus, as shown for the *KpnI* 1.2 kb sequences in Fig. 1, lanes *a-f*. No hybridization with the cytoplasmic poly(A)⁺ or poly(A)⁻ RNA fractions was observed with the various probes, even with ~30-fold longer exposure times. In contrast, *Alu* family sequences could be readily detected in these experimental conditions in the nuclear poly(A)⁺ and poly(A)⁻ RNA fractions and in the cytoplasmic poly(A)⁺ fraction, although they did appear to be more abundant in the nuclear poly(A)⁻ RNA.

The use of a cDNA probe for the 3' end of a human β -actin gene probe as a control (Fig. 1, lanes *j*, *l*) showed that the heterogeneous lengths observed for the *KpnI* transcripts could not be the result of RNA degradation during isolation. However, the possibility remains that the heterogeneity may arise from the more or less random *in vivo* processing of transcripts of discrete lengths.

The *in vivo* synthesis of RNA homologous to the *KpnI* family sequences was studied by growing HeLa cells in the presence of ³²P-orthophosphate. The labelled nuclear RNA was then isolated and analysed by either hybridization to Southern blots of the *KpnI* plasmids (Fig. 2A), or by first purifying the RNA homologous to a specific plasmid and then subjecting it to polyacrylamide gel electrophoresis. Figure 2A shows that RNA transcripts homologous to *KpnI* 1.2-1.9 kb sequences are synthesized in HeLa nuclei (lanes *a-f*). No hybridization was observed with the *KpnI* 1.8 kb sequences homologous to pBK(1.8)20 (lane *d*). In confirmation of the results in Fig. 1, Fig. 2B shows that the *in vivo* labelled *KpnI* transcripts are heterogeneous in length. As previously reported¹⁴, some high molecular weight RNAs bound nonspecifically to the filters (Fig. 2B, lanes *e*, *f*): this RNA was removed by treating the filters with T₁ ribonuclease, whereas RNA specifically bound to the *KpnI* sequences was relatively unaffected by T₁ nuclease treatment (data not shown).

The specific RNA polymerase involved in the transcription of the *KpnI* sequences was determined using an *in vitro* nuclear transcription system in which the incorporation of [α -³²P]UTP into RNA occurs in the presence of various concentrations of α -amanitin¹⁵. RNA polymerases II and III are readily distinguished by their sensitivities to α -amanitin: greater than 80% of RNA polymerase II activity is eliminated in the presence of 0.6 µg ml⁻¹ of α -amanitin, whereas polymerase III is active *in vitro* up to concentrations of about 200 µg ml⁻¹ α -amanitin. In these experiments, cloned inserts representing six different

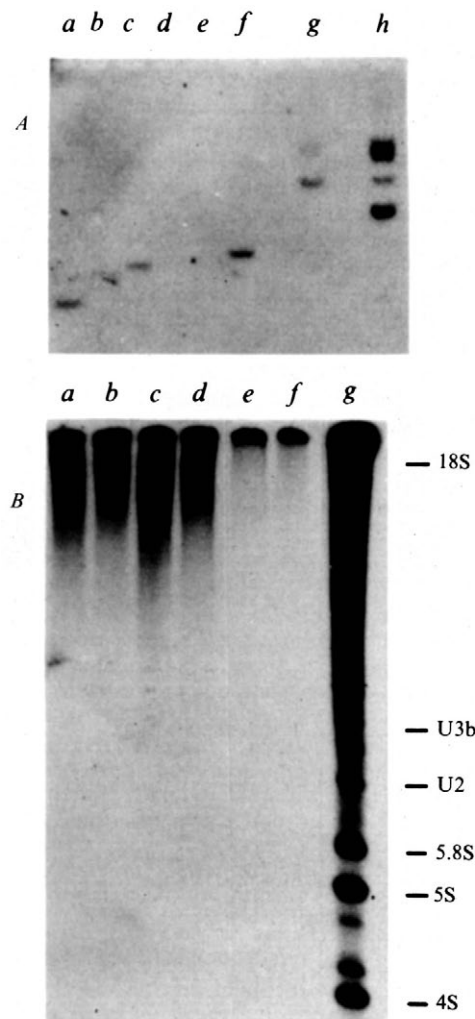


Fig. 2 Hybridization of ^{32}P -labelled nuclear RNA to cloned *KpnI* DNA segments bound to nitrocellulose filters. **A**, Hybridization of ^{32}P -*in vivo* labelled RNA (5×10^6 c.p.m.) to a Southern blot containing the inserts of the various *KpnI* plasmids as follows: a, pBK(1.2)11; b, pBK(1.5)54; c, pBK(1.8)26; d, pBK(1.8)20; e, pBK(1.8)11; f, pBK(1.9)51. Lane g contained a partial *EcoRI* digest of the human β -actin cDNA clone. Lane h contained an *EcoRI* digest of the human β -actin cDNA clone. **B**, Polyacrylamide gel electrophoresis of *in vivo* labelled RNAs eluted from filters containing the various *KpnI* plasmids as follows: lane a, pBK(1.2)11; b, pBK(1.5)54; c, pBK(1.8)26; d, pBK(1.8)20; e, pBK(1.8)11; f, pBK(1.9)51. Lane g contained a partial *EcoRI* digest of the human β -actin cDNA clone. Lane h contained an *EcoRI* digest of the human β -actin cDNA clone.

Methods: HeLa cells at $2 \times 10^6 \text{ ml}^{-1}$ were labelled for 16 h by growth in phosphate-free Eagle's medium containing 10% (v/v) dialysed fetal calf serum and 0.3 mCi ml^{-1} of ^{32}P -orthophosphate. The nuclear RNA, which was labelled to a specific activity of about 10^6 c.p.m. μg^{-1} , was then isolated as described in the legend to Fig. 1. In **A**, the same filter as that used in the hybridization with *in vitro* labelled RNA (Fig. 3A) was used. The probe was removed by successive washes of 20 min each in $0.1 \times \text{SSC}$, 0.1% SDS, 1% Na-pyrophosphate at 85°C until no autoradiographic signal could be detected. Electrophoresis and blotting conditions were as previously described⁸. Hybridization was carried out essentially according to Southern²⁴ in $5 \times \text{SSC}$, 50 mM sodium-phosphate, $\text{pH } 6.8$, 0.5% bovine serum albumin, 0.5% polyvinylpyrrolidone 360, 0.5% Ficoll 400 and $50 \mu\text{g ml}^{-1}$ *Escherichia coli* tRNA at 65°C for 48 h. The filters were washed and exposed to film as described in the legend to Fig. 1. In **B**, $20 \mu\text{g}$ of each of the plasmids containing *KpnI* family sequences were immobilized on 1 cm^2 nitrocellulose filters¹⁴ and hybridized to 2×10^6 c.p.m. of RNA labelled *in vivo* as described above. Hybridization and washing conditions were as described in the legend to Fig. 1. It should be noted that under these stringent conditions, no more than 0.75% of the total input radioactivity was recovered as RNA bound to the filters. The labelled RNA was then eluted from the filters at 85°C for 5 min in 1 ml distilled water previously treated with 0.1% (v/v) diethylpyrocarbonate. The eluted RNA was ethanol precipitated in the presence of $10 \mu\text{g ml}^{-1}$ yeast tRNA and resuspended in $25 \mu\text{l}$ of 90% (v/v) formamide, 89 mM Tris, 89 mM boric acid, 2 mM EDTA, $\text{pH } 8.7$. The RNA was heated for 1 min at 65°C and electrophoresed in a 5% (w/v) polyacrylamide- 7 M urea gel in the Tris-borate-EDTA buffer²¹. The gel was dried and exposed to film for 2 weeks as described in the legend to Fig. 1.

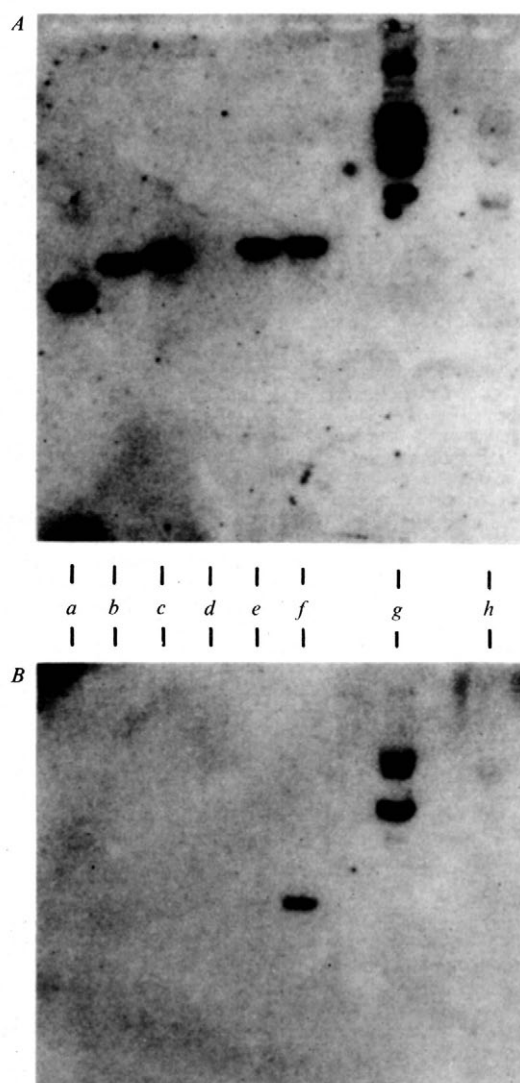


Fig. 3 Hybridization to cloned human *KpnI* DNA segments of labelled nuclear RNA synthesized *in vitro* in the presence and absence of α -amanitin. **A** Shows the autoradiograms obtained with labelled RNA synthesized in the absence of α -amanitin (1.2×10^8 total c.p.m.). **B** Shows the results obtained in the presence of $0.6 \mu\text{g ml}^{-1}$ of α -amanitin (1.5×10^7 total c.p.m.). Times of exposure were corrected for the differences in total counts added to the hybridization mixtures which contained constant cell equivalents of RNA. Lanes a-f contained inserts from the various *KpnI* plasmids as follows: a, pBK(1.2)11; b, pBK(1.5)54; c, pBK(1.8)26; d, pBK(1.8)20; e, pBK(1.8)11; f, pBK(1.9)51. Lane g contained a partial *EcoRI* digest of the *Alu* probe Blur 8 and lane h contained an *EcoRI* digest of the human β -actin cDNA clone.

Methods: Isolated inserts from plasmids containing cloned human *KpnI* sequences were resolved in a 0.8% agarose gel and the DNA transferred to nitrocellulose²⁴. HeLa nuclei were isolated as described in Fig. 1 and suspended in 25% glycerol, 5 mM Mg-acetate, 0.1 mM EDTA, 10 mM β -mercaptoethanol and 50 mM Tris-HCl, $\text{pH } 7.9$. For *in vitro* transcription, $275 \mu\text{l}$ of nuclei (from $\sim 6 \times 10^7$ cells) were incubated in each $500 \mu\text{l}$ reaction mixture containing 13.5% glycerol, 60 mM Tris-HCl, $\text{pH } 7.9$, 6 mM Mg-acetate, 1 mM MnCl_2 , 150 mM KCl, $50 \mu\text{M}$ EDTA, 15 mM β -mercaptoethanol, 1.2 mM ATP, 0.75 mM CTP and GTP, and $250 \mu\text{Ci}$ of $[\alpha\text{-}^{32}\text{P}] \text{UTP}$ ($\sim 2,500 \text{ Ci mmol}^{-1}$)²⁵. Where indicated, $0.6 \mu\text{g ml}^{-1}$ of α -amanitin was included in the reaction mixture. After 60 min, an equal volume of acetate buffer (0.2 M Na-acetate, 2 mM EDTA, $\text{pH } 5.5$) was added. SDS was added to a final concentration of 0.5% and immediately thereafter the mixture was extracted with phenol at 65°C . RNA was further purified by trichloroacetic acid (TCA) precipitation as described by Groudine *et al.*²⁶. The dried ethanol precipitates were resuspended in $500 \mu\text{l}$ of 30 mM Na-pyrophosphate and precipitated by the addition of an equal volume of 20% TCA. The precipitates were collected on Millipore nitrocellulose filter disks and washed thrice with 5% TCA, 30 mM Na-pyrophosphate. The filters were transferred to RNase-free baked glass vials and the RNA eluted with 10 mM Tris-HCl, $\text{pH } 7.5$, 5 mM EDTA and 1% SDS at 65°C . ^{32}P -RNAs in the eluates were collected by ethanol precipitation and used as probes in the blot hybridizations. Hybridization, filter washing and film exposure were carried out as described in the legend to Fig. 1 except that the hybridization mixture contained $200 \mu\text{l ml}^{-1}$ of yeast tRNA and hybridization at 65°C was for 72 h.

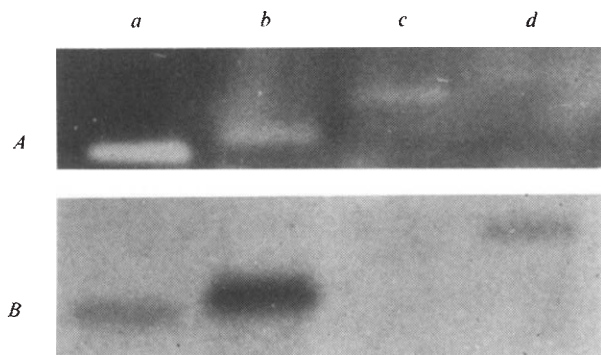


Fig. 4 Strand separation of *KpnI* 1.2 and 1.5 kb DNA segments and hybridization with ^{32}P -labelled nuclear RNA synthesized *in vivo*. **A** Shows a stained gel in which the two complementary strands of the inserts of two of the *KpnI* plasmids, purified as described below, have been run in separate lanes. Lanes *a* and *b* contain the fast and slow strands, respectively, of the pBK(1.2)11 insert, and lanes *c* and *d* contain the fast and slow strands of the pBK(1.5)54 insert. **B** Shows the corresponding autoradiograms, obtained after hybridization to *in vivo* labelled nuclear RNA.

Methods: 11 μg of the plasmids pBK(1.2)11 and pBK(1.5)54 were restricted with *KpnI*, ethanol precipitated and resuspended in 100 μl of a solution containing 30% (v/v) dimethyl sulphoxide, 1 mM EDTA, 0.025% xylene cyanol and 0.025% bromophenol blue. The samples were heated at 97 $^{\circ}\text{C}$ for 3 min, chilled in ice-water, and electrophoresed at 6 V cm^{-1} for 4 h in a 1% agarose gel in 50 mM Tris-HCl, pH 8.4, 0.38 M glycine, and 2 mM EDTA²⁷. Agarose strips containing the separated strands were frozen in an ethanol-dry-ice bath and then thawed at 55 $^{\circ}\text{C}$ for 10 min. This procedure was repeated five times. The agarose slurry was centrifuged at 3,000g for 10 min and the supernatant removed and saved. The pellets were extracted twice with 0.5 M NH_4 -acetate and the combined supernatants then adjusted to 0.5 M with NH_4 -acetate. DNA was then ethanol precipitated in the presence of 10 mM MgCl_2 and 10 $\mu\text{g ml}^{-1}$ of purified yeast tRNA. The pellets were vacuum dried, resuspended in water containing 15% sucrose, 0.4% orange G, 2.4 mM EDTA and 0.02% SDS. The complementary strands were electrophoresed in separate lanes in a 1% agarose-TBE buffer gel system (Fig. 2). The DNA was transferred to dibenzoyloxymethyl (DBM) paper in 1 M Na-acetate as described by Christophe *et al.*²⁸. After transfer, the DBM filter was washed once in 0.5 N NaOH and twice in $2\times\text{SSC}$, each for 30 min at room temperature. The filters were hybridized with 10^7 c.p.m. of ^{32}P -labelled nuclear RNA labelled *in vivo* as described in the legend to Fig. 2. Hybridization conditions, washes and exposure times were as described in Fig. 3.

KpnI family members were immobilized on nitrocellulose filters and hybridized with ^{32}P -labelled RNA synthesized *in vitro* in the presence or absence of α -amanitin. A human *Alu* sequence cloned in the *Bam*HI site of pBR322 (Blur 8) and a human β -actin cDNA sequence served as controls. Figure 3A shows that in the absence of α -amanitin, transcripts of the *Alu* sequences (lane *g*), the β -actin gene sequences (lane *h*) and all but one (that represented by plasmid pBK(1.8)20, lane *d*) of the *KpnI* families (lanes *a*–*f*) were synthesized.

Figure 3B shows that 0.6 $\mu\text{g ml}^{-1}$ α -amanitin inhibited transcription of all but one of the *KpnI* sequences, and of the β -actin gene. *Alu* sequences and sequences homologous to pBK(1.9)51, however, continued to be transcribed, although at a reduced rate. Plasmid pBK(1.9)51 actually contains an *Alu*-like sequence⁸, so it was not possible to determine which of the RNA polymerases was responsible for the transcription of the *KpnI* sequences in this plasmid. 200 $\mu\text{g ml}^{-1}$ of α -amanitin inhibited transcription of all the *KpnI* sequences and the *Alu* sequences. Thus, unlike the *Alu* family sequences, which are transcribed extensively by RNA polymerase III¹⁶, the *KpnI* family sequences are transcribed by RNA polymerase II. These results suggest that either polymerase II promoters exist in *KpnI* DNAs or that their transcription occurs by extension of RNA polymerase II transcripts initiating upstream from the sequences. In the latter case, *KpnI* repeats must be close to these polymerase II transcription units since nascent chains extend only about 500 bases in the *in vitro* system¹⁷.

That the polymerase transcribes the *KpnI* sequences asymmetrically was indicated by the following experiment. The separated, complementary strands of the inserts in pBK(1.2)11 and in pBK(1.5)54 were hybridized to HeLa nuclear RNA

labelled *in vivo* with ^{32}P -orthophosphate. As shown in Fig. 4, the labelled RNA hybridized predominantly with the slow strands of both inserts. Experiments with RNA labelled with ^{32}P -UTP in the *in vitro* system also indicated asymmetric transcription of both strands. In this case, RNA sequences homologous to the slow strands of *KpnI* 1.2 and 1.5 kb DNAs were 2.5–3.5 times more abundant than transcripts homologous to the corresponding fast strands.

We thank Drs L. Kedes and P. Ponte for the use of the human cytoskeletal β -actin cDNA clone and Ms R. Alterman and Dr A. Skoultschi for their help. This work was supported by NIH grant CA16790 (NCI) and Core support for the Cancer Research Center, NIH NCIP30-CA1330. B.S.-Z. is supported by NIH Research Training in Medical Genetics, no. 5T32GM07001, and P.J.Z. is supported by National Research Award 5T32 CA09060-08.

Received 16 January; accepted 18 May 1983.

- Davidson, E. H. & Britten, R. J. *Science* **204**, 1052–1059 (1979).
- Schmid, C. W. & Jelinek, W. R. *Science* **216**, 1065–1070 (1982).
- Musich, P. R., Brown, F. L. & Maio, J. J. *Chromosoma* **80**, 331–348 (1980).
- Adams, J. W., Kaufman, R. E., Kretschmer, P., Harrison, M. & Nienhuis, A. W. *Nucleic Acids Res.* **8**, 6113–6128 (1980).
- Schmeckpeper, B. J., Willard, H. F. & Smith, D. D. *Nucleic Acids Res.* **9**, 1853–1872 (1981).
- Maio, J. J., Brown, F. L., McKenna, G. W. & Musich, P. R. *Chromosoma* **83**, 127–144 (1981).
- Jagadeeswaran, P. *et al.* *ICN-UCLA Symp. molec. Cell Biol.* **23**, 71–84 (1981).
- Shafit-Zagardo, B., Maio, J. J. & Brown, F. L. *Nucleic Acids Res.* **10**, 3175–3193 (1982).
- Manuelidis, L. *Nucleic Acids Res.* **10**, 3211–3219 (1982).
- Manuelidis, L. & Biro, P. A. *Nucleic Acids Res.* **10**, 3221–3239 (1982).
- Shafit-Zagardo, B., Brown, F. L., Maio, J. J. & Adams, J. W. *Gene* **20**, 397–407 (1982).
- Gillespie, D., Adams, J. W., Costanzi, C. & Caranfa, M. J. *Gene* **20**, 409–414 (1982).
- Grimaldi, G. & Singer, M. F. *Nucleic Acids Res.* **11**, 321–338 (1983).
- Weiner, A. M. *Cell* **22**, 209–218 (1980).
- Roeder, R. G. in *RNA Polymerase* (eds Losick, R. & Chamberlin, M.) 285–329 (Cold Spring Harbor Laboratory, New York, 1976).
- Duncan, C. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **76**, 5095–5099 (1979).
- Hofer, E. & Darnell, J. E. *Cell* **23**, 585–593 (1981).
- Berger, S. L. & Birkmeier, C. S. *Biochemistry* **18**, 5143–5149 (1979).
- Soeiro, R. & Darnell, J. E. *J. molec. Biol.* **44**, 551–562 (1969).
- Nevins, J. R. *Meth. Enzym.* **65**, 768–785 (1980).
- Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning* (Cold Spring Harbor Laboratory, New York, 1982).
- Thomas, P. S. *Proc. natn. Acad. Sci. U.S.A.* **72**, 5201–5205 (1980).
- Wahl, G. M., Stern, M. & Stark, G. R. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3683–3687 (1979).
- Southern, E. M. *J. molec. Biol.* **93**, 503–517 (1975).
- McKnight, G. S. & Palmiter, R. D. *J. biol. Chem.* **254**, 9050–9058 (1979).
- Groudine, M., Peretz, M. & Weintraub, H. *Molec. cell. Biol.* **1**, 281–288 (1981).
- Maxam, A. M. & Gilbert, W. *Meth. Enzym.* **65**, 499–580 (1980).
- Christophe, D., Brocas, H. & Vassart, G. *Analyt. Biochem.* **120**, 259–261 (1982).

DNA sequences at the ends of transposon Tn5 required for transposition

Reid C. Johnson & William S. Reznikoff

Department of Biochemistry, College of Agricultural and Life Sciences, University of Wisconsin-Madison, Madison, Wisconsin 53706, USA

Transposons are a class of genetic elements that can move from one site in a cell's genome to another independently of the cell's general recombination system. Little is known about the mechanism of transposition of compound transposons such as Tn5, but it is thought that a transposon-encoded protein (a transposase) must recognize the outer ends of the element and, together with host factors, catalyse the transfer of the internal DNA into a new site in a manner that may involve replication. It has previously been shown that the synthesis of an IS50R-encoded protein (protein 1) is an essential requirement for Tn5 transposition^{1–5}. Here we demonstrate that a structure containing only the outer 186 base pairs (bp) of both inverted repeats is capable of being efficiently complemented to transpose in *Escherichia coli*, provided IS50R is located close by on the same replicon. In addition, *Bal*31-generated deletions indicate that 16–18 bp of the outer end of IS50L are required for

transposition. This 16–18-bp sequence contains the 8–9-bp small inverted repeat present at each end of IS50 plus a 9-bp sequence which is homologous to an interrelated sequence present in four copies in the chromosomal origin of replication in a variety of Gram-negative bacteria. This sequence organization suggests that the ends of Tn5 may function to provide a recognition site for the Tn5 transposase adjacent to a sequence recognized by the host replication system.

Tn5 is an example of a composite prokaryotic transposable element. It contains two physically similar but genetically different 1,534-bp insertion sequences (IS50R and IS50L) in inverted orientation flanking a 2.7-kilobase (kb) segment of unique DNA which contains a gene encoding an aminoglycoside antibiotic modifying function^{6–9}. A derivative of Tn5 was constructed containing the outer 186 bp of the element in inverted orientation flanking a 2-kb segment of DNA encoding tetracycline resistance. While this structure is unable to transpose in the absence of a functional copy of IS50R, transposition of Tn5-320 from a ColE1 plasmid to λ was detected when a wild-type copy of Tn5 was present on an F plasmid³. This complementation was very inefficient, as Tn5-320 only transposed at 1–3% the frequency of the wild-type element. Table 1 shows that complementation remains inefficient when the donor Tn5 is present on a multi-copy plasmid. However, when the wild-type donor is located on the same replicon within 420 bp of Tn5-320, both elements transpose at approximately equal frequencies. This result indicates that Tn5-320 contains all the DNA sequence information required for efficient translocation when the transposase is synthesized in close proximity. The inability of the transposase to act efficiently in *trans* has been demonstrated in several studies with Tn5 (refs 2–4) and other Tn5-like transposable elements (for example, Tn10, Tn903, IS1)^{10–13} and may be due to an unstable nature of the active form of the transposase or a high affinity for DNA at the site of synthesis followed by a limited one-dimensional diffusion along the DNA¹³.

To investigate the sequence requirements for transposition further, we varied the amount of sequence present in the left inverted repeat (IS50L) while leaving the functional right repeat (IS50R) intact. Table 2 shows that efficient transposition occurs in derivatives such as Tn5-341 and Tn5-346 which contain only the outer 56 bp of the left inverted repeat. In addition, the sequence immediately adjacent to the left repeat, including the 9 bp of duplicated host DNA⁸, which is different in Tn5-341 and Tn5-346, has little effect on subsequent transposition.

pRZ341 was constructed so that a unique *Bam*HI restriction site was located at nucleotide 53 from the left end of the

```

R1  TTATCCACA
R2  TTATACACA
R3  TTATCCAAA
R4  TTATCCACA*
Tn5 TTATACACA

```

Fig. 1 Homology at the ends of Tn5 with a repeated sequence in the chromosomal origin of replication (*oriC*). The Tn5 sequence from nucleotides 8–16 is given with the four copies (R1–R4) of the interrelated sequence in the *oriC* region of six different Gram-negative bacteria^{14,15}. A large letter indicates that the nucleotide is totally conserved; a small letter indicates it is the predominant nucleotide. C* indicates the location of a single base pair change in the *E. coli oriC* which inactivates its function²⁰.

element. By linearizing pRZ341 at the *Bam*HI site, treating with *Bal*31 exonuclease, and re-ligating in the presence of *Bam*HI107ntaining oligonucleotides (see Table 3), we were able to remove progressively the remaining sequences in the left inverted repeat. Subsequent transposition assays revealed that deletions up to nucleotide 21 (Tn5-341:21 actually contains the outer 23 bp of Tn5 as the first 2 bp of the *Bam*HI linker is equivalent to nucleotides 22 and 23 of Tn5) had no effect on the transposition frequency (Table 3). An element (Tn5-341:18) containing only the outer 18 bp is still able to transpose although at a reduced frequency. However, removal of an additional 3 or more bp (Tn5-341:15, 13 or 7) results in a drastic lowering of the transposition proficiency of the element to the point where it is undetectable in our assays.

These results indicate that the outer 16–18 bp of DNA are essential for efficient transposition, although maximum transposition may require sequences up to nucleotide 23. The outer 9 bp of IS50 are repeated with one mismatch in inverted orientation at the inner end of the insertion sequence^{8,9}, which has led to the proposal that this sequence may constitute at least part of the recognition determinant for the IS50R-encoded transposase^{13,21}. It is perhaps surprising that Tn5-341:15 and Tn5-341:13, which contain the terminal 9 bp sequence, are unable to transpose since both IS50 or a dimeric Tn5 containing a direct repeat structure transpose at relatively high efficiency^{9,21}. This indicates that an inner end of IS50 can substitute for one outer end to mediate transposition even though it contains only the terminal 8/9 bp in common with the outer end. We have not yet investigated the sequence requirements of an IS50 inner end nor do we know whether the transposition process of IS50-like elements is mechanistically identical to a normal Tn5 structure.

The sequences adjacent to the terminal 9 bp repeat (nucleotides 7–16) display remarkable similarity to a highly homologous 9-bp sequence which is repeated four times within the *oriC* region of six different Gram-negative bacteria (Fig. 1)^{14,15}. The significance of this homology is suggested by the difference in the transposition proficiency of Tn5-341:18 which leaves this sequence (plus an additional 2 bp) intact, contrasting with the transposition-defective Tn5-341:15 in which the last base pair of this sequence is absent. The *oriC* homologous sequence is not present at the inner end of IS50. However, Isberg *et al.*² and Berg and co-workers^{9,21} have concluded that the outer and inner ends of IS50 are not functionally symmetrical. Isberg *et al.*² determined that a pair of inner ends mediated transposition at <0.2% the frequency of the outer ends. In contrast, similar studies have shown that the inner and outer ends of the structurally related transposon Tn10 do function with equal efficiency¹⁰.

The *oriC*-like sequence found at the outer ends of Tn5 may serve to direct host replication functions to the ends of the element during transposition. The requirement for this sequence at the end distal to the site of transposase synthesis suggests that transposition of Tn5 may involve replication

Table 1 Complementation of Tn5-320

Exp	Tn5 elements present	Transposition frequency		
		Tet ^r	Kan ^r	Tet ^r /Kan ^r
A	ColE1::Tn5-320	<6.6×10 ⁻⁹	—	—
	ColE1::Tn5-320			
B	+	4.6×10 ⁻⁷	1.1×10 ⁻⁵	0.04
	RSF1010::Tn5-wt			
C	ColE1::Tn5-320:	4.8×10 ⁻⁶	5.2×10 ⁻⁶	0.92
	Tn5-wt			

Cultures of RZ502Δbhn (*E. coli* W3110 *lac recA56 srl str λcI857 b515 b519 nin5 Sam 7*)³ containing the various plasmids were heat induced and the resulting lysates used to lysogenize Hf1–1 as described previously¹. The transposition frequency measures the kanamycin (Kan = Tn5-wild-type, wt) or tetracycline (Ter = Tn5-320) resistant transductants per plaque-forming unit of the phage. Expt A, no IS50R functions are supplied; expt B, IS50R functions are supplied on the multi-copy compatible plasmid RSF1010; expt C, wild-type Tn5 was transposed onto ColE1::Tn5-320 and is located ~420 bp from Tn5-320 with IS50R oriented adjacent to Tn5-320. In expt C, 3% of the Kan^r transductants were also Tet^r and 11% of the Tet^r transductants were also Kan^r due to transposition of Tn5-320::Tn5-wild-type. The doubly resistant transductants were not included in the values shown. Tn5-320 was constructed as elsewhere described³. Wild-type Tn5 was transposed onto RSF1010 by Jerry Yin.

Table 2 Transposition assays of Tn5 derivatives containing various amounts of the left inverted repeat

Plasmid present	Structure of Tn5	Left end sequence present	Transposition frequency
pRZ190		1-1,534	9.4×10^{-6}
pRZ191		1-118	1.1×10^{-5}
pRZ341		1-56	1.8×10^{-5}
pRZ346		1-56	1.8×10^{-5}
pRZ340		0	$< 5.0 \times 10^{-9}$

Transposition assays of Tn5 derivatives containing various amounts of the left inverted repeat. Transposition assays were performed in RZ102Δbnn (RZ102 = *E. coli* MO *recA56 str*)³ as described in Table 1 legend. pRZ190 is ColE1-Tn5 containing a 2.7-kb *Bgl*III restriction fragment which carries the tetracycline resistance determinant inserted into the *Bam*HI site in Tn5 (S. J. Rothstein and W.S.R., unpublished). pRZ191 has the DNA from the *Sal*I site in pRZ190 to the *Hae*III site in the left inverted repeat deleted⁵. pRZ341 was constructed by ligating a 120-bp *Eco*RI-*Sau*3a restriction fragment, which contains the outer 56 bp of IS50L plus ColE1 sequences, from pRZ102⁷ into pBR322 between the *Eco*RI site and the *Bam*HI site, regenerating the *Bam*HI restriction site. An *Eco*RI-*Sal*I fragment from this plasmid was then substituted for the left inverted repeat of pRZ190. pRZ346 was made in an analogous manner starting with a Tn5 insertion into pBR322 with IS50R located in the tetracycline resistance region about 45 bp from the *Eco*RI site. Therefore, the left end of Tn5-346 is actually derived from IS50R and is adjacent to pBR322 sequences. pRZ340 contains non-Tn5 DNA at its left end and was constructed by ligating an *Eco*RI-*Sal*I fragment from pRZ4006 (a pBR322 derived plasmid containing the *lacPO* between the *Eco*RI and *Bam*HI sites) into pRZ190.

Table 3 Transposition assays on *Bal*31 generated deletions into the left inverted repeat of Tn5

1	10	20	30	Tn5 mutant	Transposition frequency
CTGACTCTTATACACAAGTAGCGTCTTGAA				Tn5-341	1.7×10^{-5}
.....CGGGA				Tn5-341:25	1.3×10^{-5}
.....CGGGATCCC				Tn5-341:21	1.7×10^{-5}
.....CGGGATCCCCG				Tn5-341:18	2.8×10^{-6}
.....CGGGATCCCCG				Tn5-341:15	$< 1.0 \times 10^{-9}$
.....CGGGATCCCCG				Tn5-341:13	$< 1.0 \times 10^{-9}$
.....CGGGATCCCCG				Tn5-341:7	$< 1.0 \times 10^{-9}$

Transposition assays on *Bal*31-generated deletions into the left inverted repeat of Tn5. The nucleotide sequence of the outer 30 bp of Tn5 is given along with the location of the *Bam*HI containing oligonucleotide (in italics) present in the deletion mutants. A dot indicates the wild-type nucleotide is present. The arrow depicts those bases which are repeated in inverted orientation at the inner end of IS50 and the underlined sequence contains the nucleotides which are homologous to a conserved sequence found in four copies in the *oriC* region (see Fig. 1). The transposition assays were performed in RZ201Δbnn as described in the Table 1 legend. The deletion mutants were constructed as follows. pRZ341 (Table 2) was linearized at its *Bam*HI restriction site and subjected to *Bal*31 exonuclease (New England Biolabs) digestion at 1 U ml⁻¹ in a 100-μl reaction volume containing 0.2 M NaCl, 12 mM CaCl₂, 12 mM MgCl₂, 1 mM EDTA, 20 mM Tris-HCl, pH 7.9, and 2.4 pmol of DNA termini. The reaction was terminated after 30–60 s at 30 °C by the addition of diethyl pyrocarbonate and EDTA to 0.05% and 30 mM, respectively. After ethanol precipitation, ~0.2 pmol of termini were ligated overnight at 15 °C in a 30–50 μl reaction volume with 24 pmol of *Bam*HI containing oligonucleotides (New England Biolabs) which were phosphorylated using polynucleotide kinase and ATP. The reaction mixture was then heated to 65 °C for 10 min to inactivate any remaining ligase and digested with 30 U of *Bam*HI for 4 h. The reaction was again heated to 65 °C for 10 min and the plasmid DNA containing *Bam*HI sticky ends was circularized in the presence of T4 DNA ligase (gift of R. Simoni) at a plasmid terminus concentration of approximately 1 nM and transformed into RZ211 (Δ(lac-pro) *recA56 srl str*). The exact end points of the deletions were determined using the DNA sequencing methods of Maxam and Gilbert¹⁹.

proceeding from both ends in a symmetrical manner. This type of mechanism may differ from the mechanism of transposition of other elements which is believed to involve polarized replication beginning at the end proximal to the site of transposase synthesis (for example, Tn9 (ref. 16)). Perhaps this difference is reflected in the very low frequency of co-integration events promoted by Tn5 in the absence of homologous recombination, in contrast to many other transposons (including Tn9)^{4,17,18}. We have searched the outer 100 bp of several other transposable elements, including Tn10, IS1/Tn9, Tn3 and bacteriophage Mu, and have found no significant homology to the *oriC* repeated sequence.

We thank L. Munson for initially pointing out the homology to *oriC* and J. Yin and S. Stibitz for valuable discussion. R.C.J. was supported in part by a NIH training grant (GM07215). Additional support was from grants from the NSF (PCM791086) and the NIH (GM19670) to W.S.R.

Received 20 April; accepted 17 May 1983.

- Rothstein, S. J., Jorgensen, R. A., Postle, K. & Reznikoff, W. S. *Cell* **19**, 795–805 (1980).
- Isberg, R. R., Lazaar, A. L. & Syvanen, M. *Cell* **30**, 883–892 (1982).

- Johnson, R. C., Yin, J. C.-P. & Reznikoff, W. S. *Cell* **30**, 873–882 (1982).
- Isberg, R. R. & Syvanen, M. *J. molec. Biol.* **150**, 15–32 (1981).
- Rothstein, S. J. *et al. Cold Spring Harb. Symp. quant. Biol.* **45**, 99–105 (1981).
- Berg, D. E., Davies, J., Allet, B. & Roehaia, J. D. *Proc. natn. Acad. Sci. U.S.A.* **72**, 3628–3632 (1975).
- Jorgensen, R. A., Rothstein, S. J. & Reznikoff, W. S. *Molec. gen. Genet.* **177**, 6047–6051 (1979).
- Auerswald, E. A., Ludwig, G. S. & Schaller, H. *Cold Spring Harb. Symp. quant. Biol.* **45**, 107–114 (1981).
- Berg, D. E., Johnsrud, L., McDivitt, L., Ramabhadran, R. & Hirschel, B. J. *Proc. natn. Acad. Sci. U.S.A.* **79**, 2632–2635 (1982).
- Foster, T., Davis, M. A., Takeshita, K., Roberts, D. E. & Kleckner, N. *Cell* **23**, 201–213 (1981).
- Grindley, N. D. F. & Joyce, C. M. *Cold Spring Harb. Symp. quant. Biol.* **45**, 125–132 (1981).
- Machida, Y., Machida, C., Ohtsubo, H. & Ohtsubo, E. *Proc. natn. Acad. Sci. U.S.A.* **79**, 277–281 (1982).
- Morimoto, D., Way, J. C., Kim, H.-H. & Kleckner, N. *Cell* **32**, 799–807 (1983).
- Zyskind, J. W., Harding, N. E., Takeda, Y., Cleary, J. M. & Smith, D. W. *ICN-UCLA Symp. molec. cell. Biol.* **22**, 13–25 (1981).
- Zyskind, J. W., Cleary, J. M., Brusilow, W. S. A., Harding, N. E. & Smith, D. W. *Proc. natn. Acad. Sci. U.S.A.* **80**, 1164–1168 (1983).
- Chandler, M. J. & Galas, D. J. *J. molec. Biol.* (in the press).
- Hirschel, B. J., Galas, D. J. & Chandler, M. *Proc. natn. Acad. Sci. U.S.A.* **79**, 4530–4534 (1982).
- Berg, D. E. *Proc. natn. Acad. Sci. U.S.A.* **80**, 792–796 (1983).
- Maxam, A. M. & Gilbert, W. *Meth. Enzym.* **65**, 499–560 (1980).
- Hirota, Y., Oka, A., Sugimoto, K., Asada, K., Sasuki, H. & Takanami, M. *ICN-UCLA Symp. molec. cell. Biol.* **22**, 1–12 (1981).
- Sasakawa, C. & Berg, D. E. *J. molec. Biol.* **159**, 257–271 (1982).

Fourier and Duffieux: a French connection

P.W. Hawkes

The Fourier Transform and its Applications to Optics, 2nd Edn.

By P.M. Duffieux. Wiley: 1983. Pp.197. £32.25, \$45.25.

THE French edition was published in 1970 and was itself a revised version of the text of 1946: why should Wiley bring out an English translation of an old book by someone as obscure as P. M. Duffieux?

The answer is simple. Duffieux was one of the great original spirits of modern optics and it is entirely right that efforts should be made to rescue his name from obscurity. Virtually single-handed, he founded the subject that we now call Fourier optics and introduced the notion of spatial frequency. Moreover all this was achieved despite a woefully inadequate knowledge of the necessary mathematics.

The creation of Fourier optics revealed the intimate connection between image-forming devices and communication channels. It became apparent that optical instruments filter the information presented to them by the object under study and that this filter takes a very simple form if we study its effect on the Fourier transform or "spatial frequency spectrum" of the object. Moreover, Fourier methods have generated a new branch of coherence studies as well as providing a suitable vocabulary for discussing the information-carrying capacity of microscopes and the relation between this and their traditional resolving power.

Ironically, Duffieux first encountered the Fourier transform just after the end of the First World War, when still a student, and his first publication thus mentioned it — but fifteen years later he had completely forgotten the Fourier transform, and it was in 1934 that Charles Fabry unwittingly provoked him to return to it: "Cherchez quelque chose qui fasse intervenir toute la période", said Fabry and in 1970 Duffieux had not forgotten: "Je lui garde, pour cette remarque, une reconnaissance et une estime que rien de ce qui nous a séparés depuis n'a pu éteindre".

The rest of the story is quite as dramatic. Duffieux found a solution to the problem posed by Fabry but no one could understand it. Duffieux, depressed, wrote his equations on the blackboard of a deserted lecture theatre, where Jean Dieudonné happened to look in and exclaimed that they were the opening terms of the Fourier series transform. A few months later, further light flooded in as a result of a chance reading of the commercial brochure describing a new (and inexpensive) harmonic analyser. "J'eus une révélation hallucinante", Duffieux tells us, which was effectively the practical meaning of the

Fourier (series) transform.

In 1941, Duffieux presented his ideas to the Société de Physique but again no one understood. Write a book about it, urged Pierre Fleury, and in 1946 or 1947 the Rennes printing firm of Oberthur delivered *L'Intégrale de Fourier et ses Applications à l'Optique* to its author, by then a professor at the University of Besançon, who published it privately. To obtain a copy,



Duffieux: "J'eus une révélation hallucinante". one used to write to the professor, who, to judge from the wrapping and the handwritten label, despatched it himself. (Owing to the post-war paper shortage, the book contained only part of the author's manuscript; the remainder appeared in the *Annales de Physique*.)

For a decade or more, the book had little impact ("en France", Aimé Cotton told Duffieux, "vous mettez dix ans pour réussir . . . à condition que vous reveniez d'Angleterre"). It figures fleetingly in Linfoot's *Recent Advances in Optics* (1955), in connection with pupil functions, but astonishingly is not mentioned at all in his *Qualitätsbewertung Optischer Bilder* (1960) or its English version (1964).

In 1959, however, the situation changed dramatically. In Born and Wolf's *Principles of Optics* Duffieux is given proper credit, perhaps for the first time in a major treatise: "These methods [relating to imaging of extended objects] were developed chiefly by Duffieux . . . and were later extended and applied to particular problems by many writers". There was no longer any excuse for the optical community to plead ignorance of Duffieux's work. Admittedly, he is not always accorded more than a passing

reference; but in 1967 alone Röhler in his *Informationstheorie in der Optik* was referring to the convolution relation between object and image as the Duffieux Integral, Stroke (*Handbuch der Physik*, Vol. 29) repeatedly cited Duffieux's work, and Rosenhauer too mentioned him in Van Heel's *Advanced Optical Techniques*. Most subsequent texts dealing with image formation refer to him, notably Papoulis (*Systems and Transforms with Applications in Optics*, 1968) and Menzel, Mirandé and Weingärtner (*Fourier-Optik und Holographie*, 1974), in which a section is entitled "Das Duffieux-Integral". The optical world was gradually waking up to the realization that an original spirit had been overlooked.

In 1970, Masson and the CNRS published a revised edition of Duffieux's book, upon which the present translation is based. This English edition lacks some of the historical charm of the 1946 book, in which Duffieux's delight and astonishment at what he had found come through vividly, but has compensating merits, especially increased maturity with no loss of verve (as instanced by the presence of the elephant Jumbo in a section on resolution). The translation is respectable, though much of the fluency of the original is lost and the tone of some of the more personal passages has not always been caught. Thus in Duffieux's acknowledgments

Je remercie . . . d'abord, en bloc, tous les travailleurs du Laboratoire d'Optique de Besançon, depuis Jean Viénot, titulaire, jusqu'à Mme Reine, qui tient notre ménage, tous: techniciens, mécaniciens, étudiants débutants ou entrés dans la hiérarchie

is rendered

I thank . . . first of all, the people at the Besançon Optical Laboratory: Professeur Vienot [sic], Mrs. Reine the cleaning lady technicians, machinists, beginning students and new employees.

A little later "on n'écrit bien que de ce dont on a parlé" is translated horridly as "one may only write well of that which he has spoken". Despite this, most of the text reads easily and smoothly, though some of the lighter-hearted remarks are lost; thus by translating "véhicules" (p. 156 ff.) as "relays", Duffieux's "schéma optique . . . représente non pas un véhicule mais un convoi" loses its point. I was amused that "3 pieds de longueur" is translated by a "length of 1 m".

Two sections of the original have been omitted from this translation: Viénot's preface and the list of Duffieux's publications. Furthermore, a reference to Duffieux's *Le Statistique et l'Individuel* (1966) has been dropped from his general introduction. There seems no reason at all for leaving out Viénot's preface, since it sheds interesting light on the genesis of the book and indeed on the evolution of Duffieux's thinking. The failure to include the list of publications is inexcusable.

Given that the book is of largely historical interest, this simply obliges the reader to return to the French original. This was a perfect occasion to publish a complete list of the publications of P.M. Duffieux: since it has been lost, perhaps one of the optics journals could be persuaded to publish such a list together with their own review of the book.

It is very right and pleasing that Duffieux's book should be available, through a translation, to a wider audience. The loss of some of the trenchancy, even belligerence, of his prose is a small price to pay. Everyone interested in modern optics (and unable to read the original) will enjoy this English version. □

P. W. Hawkes is *Maître de Recherches* at the *Laboratoire d'Optique Electronique du CNRS, Toulouse*.

Postgate's fixation

John D. Tjepkema

The Fundamentals of Nitrogen Fixation.

By J.R. Postgate.

Cambridge University Press: 1982. Pp.252.

Hbk £20, \$37.50; pbk £7.95, \$12.95.

PRODUCTIVE agriculture depends on the use of nitrogen fertilizer or the inclusion of nitrogen-fixing species in crop rotations. Due to the rising cost of fertilizer, and the impracticality of its use under some conditions, research on nitrogen fixation has been given a high priority. This increased emphasis and breakthroughs in research methodologies have led to a spectacular increase in our knowledge of the process.

There has also been a spectacular increase in the number of books on nitrogen fixation, but many of these have been symposium volumes which rarely cover the field as whole. Thus John Postgate's book serves a very real need, being a concise, comprehensive and up-to-date review suitable for researchers, teachers or advanced students. The viewpoint is that of a microbiologist, who with his colleagues at the ARC Unit of Nitrogen

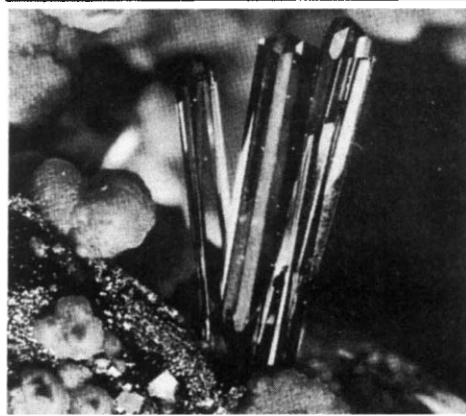
Fixation has been at the forefront of research in the area. Quite rightly it is the work of this group as well as Postgate's personal interests, such as the evolution of nitrogen fixation, that receive the greatest attention. Thus biochemistry, genetics and microbial physiology are covered in detail, while only 25 of the 200 pages of text are devoted to symbiotic nitrogen fixation. Those seeking more detailed information on this topic, which is of much current interest, or on nitrogen fixation in natural ecosystems, forestry or agriculture, are referred by Postgate to J.I. Sprent's recent book, *The Biology of Nitrogen-fixing Organisms* (McGraw-Hill, 1979), which emphasizes the more biological aspects of nitrogen fixation.

Of all the progress in nitrogen fixation research discussed by Postgate, I find that in genetics most remarkable. We started with almost no knowledge in 1970. By 1972 functioning nitrogen-fixing genes had been transferred from *Klebsiella pneumoniae* into *E.coli*, and by the end of the 1970s a detailed map of the cluster of 17 genes involved with nitrogen fixation in *K. pneumoniae* had been developed. Since the nitrogenase enzyme complex contains only three different protein subunits, this large number of genes was surprising and their study has illustrated the complexity of biological nitrogen fixation.

The style of Postgate's book is largely that of a review article, with 40 pages of references which provide a useful entry into the primary and secondary literature. Many of the sections and subsections are presented in a partially historical sequence, which gives the reader a better understanding of how present knowledge has developed. Sections vary substantially in their readability, however, and only an expert will find all of the book easy going.

Thus the book is not light reading for the lay person and would not serve very well as an introduction for the beginning student. But this is not its intended use, and overall I find the style excellent for conveying a clear picture of the current state of nitrogen fixation research to the advanced reader. □

John D. Tjepkema is in the Department of Botany and Plant Pathology at the University of Maine.



Clear crystals — the illustration (of vivianite, $\text{Fe}_3(\text{PO}_4)_2 \cdot 8\text{H}_2\text{O}$) is taken from a new wall-chart, *Minerals of the World*, published by Elsevier Scientific. The chart features 200 minerals, all in full colour; like the companion *Elsevier's Mineral and Rock Table* (for review see *Nature* 302, 183; 1983) it has been compiled by P. Lof. As well as the illustrations, the chart includes an index to the minerals shown, plus other data, and a display and general explanation of crystal forms. Price for a single copy is Dfl. 30 (\$13) and for 10 copies Dfl. 185 (\$78.75), with cheaper rates for orders of 20, 50 and 100 copies.

On the beaches of Hong Kong

C.M. Yonge

The Sea Shore Ecology of Hong Kong.

By John Morton and Brian Morton.

Hong Kong University Press: 1983.

Pp.350. Hbk \$HK 145; pbk

\$HK 120.

HONG Kong with its associated islands and the intricately zig-zagged extension of mainland territory appears as a small spot on the map of China. It may therefore come as something of a surprise to learn that islands and territory have a combined coastline one-fifth as long as that of England and Wales. The nature of these shores could hardly be more diverse, ranging from fully exposed rock pounded by great Pacific breakers through many intermediate conditions to fully sheltered mud flats. While outer shores are washed by the fully saline waters of the South China Sea, southern coasts are exposed to the outflow of the Pearl River.

So much for nature; but man has major effects, razing mountains and filling in valleys to build towns of half-million populations. Shores are built over and extended. Plover Cove, a three-mile-long arm of the sea has been cut off, pumped dry and converted into a reservoir, its original marine population replaced by freshwater plants and animals. Meanwhile pollution spreads out from the "fragrant harbour" between Hong Kong and Kowloon.

Nevertheless a diverse and fascinating population persists, adapted for intertidal life with its hazards of seasonal and diurnal fluctuations in temperature, salinity and exposure. Richer than a corresponding Atlantic flora and fauna, it is also of dual origin, a mingling of the inhabitants of more northern, Japanese waters and those of the tropical Indo-West-Pacific. This second influence is responsible for the presence of mangrove areas and of an unexpectedly large, although always subtidal, assemblage of corals. Winter exposure would be fatal.

Description of these populations is the subject matter of this book. Both authors (who are unrelated) are exceptionally well-qualified, with similar, primarily malacological, interests. The book was planned during a visit, as Royal Society Professor, of John Morton from Auckland in 1975 to Hong Kong where Brian Morton has been working since 1970. The former has unique knowledge of southern Pacific shores and is the author, with Michael Miller, of the altogether admirable *New Zealand Sea Shore* (Collins, 1968). With the aid of successive groups of students, Brian Morton has with unique energy been intensively studying the marine invertebrates of Hong Kong, adding many

times over to the substance of previous knowledge.

After a general introduction dealing with climate, hydrography and seasonal effects with zonation, four types of hard shore (including wharf piles) are described, then four kinds of soft shore with a final account of the subtidal coral shores. The book ends with a discussion of probable future problems, notably pollution and conservation on which both authors have already written extensively. In addition to 28 colour or black-and-white plates there are 132 line drawings, many occupying full

pages; both authors are capable artists and deft in the production of the pictorial diagrams of which Alan Stephenson was such a master.

This is a major addition to knowledge about the sea shore in general and also the marine fauna and flora of this region of the Pacific. Above all the book tells much about the living animal with which both authors are primarily concerned. It is a pleasure to see their work in print. □

Sir Maurice Yonge is Honorary Fellow in Zoology at the University of Edinburgh.

In the beginning . . .

W. Graham Richards

Modern Quantum Chemistry:

Introduction to Advanced Electronic Structure Theory.

By Attila Szabo and Neil S. Ostlund.
Macmillan/Collier Macmillan: 1983.
Pp. 446. \$49.95, £36.

Introduction to Advanced Electronic Structure Theory, the subtitle of Szabo and Ostlund's book, is a better indicator of the contents than the main title. Better still would have been something which emphasized that the book is largely about the *ab initio* computation of molecular wavefunctions at the highest levels of accuracy currently available. Many texts on quantum chemistry take the student to the Hartree-Fock level, but few go much further. Here, however, more than half of the content is devoted to methods of improving upon the wavefunctions which can be obtained from standard packages such as those available from the Quantum Chemistry Program Exchange.

Configuration interaction is the subject of the first of four chapters which deal with the incorporation of electron correlation. One-electron density matrices, natural orbitals, multiconfiguration self-consistent field calculations and the generalized valence bond method are covered. The size-consistency problem associated with truncating configuration interaction is used to highlight the need for many-body approaches which avoid this deficiency.

Full chapters are then devoted to pair and coupled-pair theories, to many-body perturbation theory and to the one-particle many-body Green's function. These chapters contain some fairly advanced material and despite the liberal use of examples will be hard going for many students. Probably the most receptive users of this admirable work will be researchers and graduate students who are already familiar with the computation of *ab initio* wavefunctions using programs such as the Gaussian 80 series, but who want to go beyond that level of accuracy.

In general, calculators of wavefunctions fall into two categories: those who want to do very accurate work on rather small molecules and those who are prepared to sacrifice accuracy in order to study larger molecules of more interest to experimental chemists. The authors are unashamedly in the former group and their work will be appreciated by like-minded computational chemists. In addition, one can hope that with the help of this exposition and the continuing developments in computer power, the authors' conception of accuracy will be accepted by the more pragmatic practitioners. Certainly those who merely feed data into universal programs could have their confidence troubled by learning just how much better things can be done. The description "*ab initio*" often gives a spurious hint of quality. It should not be taken as an indication of the last word in sophistication. □

*W. Graham Richards is a Lecturer in Physical Chemistry at the University of Oxford. The second edition of his *Ab Initio Molecular Orbital Calculations for Chemists* was published by Oxford University Press earlier this year.*

Biology and business

Peter Newmark

The Gene Business: Who Should Control Biotechnology?

By Edward Yoxen.
Pan: 1983. Pp. 264. Pbk £3.95.

EDWARD YOXEN's book, *The Gene Business*, is part of a series called *Crucible: Science in Society*, "designed to make critiques of the role of expert power — power through knowledge — in our lives". Each book is linked to a television documentary shown under the same series title on British television's Channel 4.

The "expert power" picked upon by Yoxen is biotechnology, and he takes the threat to the sugar industry as one of his examples. Traditionally the economic mainstay of a few underdeveloped nations, sugar now faces increasing competition from "engineered" products. Immobilized enzymes already produce high-fructose

corn syrup on a large scale. Aspartame, the combination of two amino acids produced by fermentation, promises to become widely used as a sweetener. And, he might have added, the gene for thaumatin, an exceedingly sweet component of a plant gathered from the wild in Africa, has been expressed in bacteria. In the face of those who unremittingly support the development of all such alternatives to sugar from the cane, if only because they favour the survival of teeth over that of dentistry, Yoxen waves the banner of Social Responsibility in Science and Technology.

Fortunately, he is mostly sensible about the genuinely important issues associated with the growth of biotechnology. For example he raises the question of the fate of the superseded agricultural workers if biotechnological products ever displace cane sugar, but without giving support to the far-fetched notion that American efforts in alternative sweeteners are essentially a way of attacking Cuba by driving down demand for cane sugar.

How, he asks more generally, can research be directed towards social needs and away from the twin evils of what he sees as a self-indulgent elite of research scientists and a business community preoccupied with maximum profits. His answers are few, and as often as not they seem to be reflex reactions (worker participation, public ownership). Indeed, Yoxen's "radical-reformist" policy for university-industry collaboration in biotechnology is distinctly conservative: no firms on campus, no direct university investment in the companies with which they collaborate, strict rules on academic consulting. Well before the end of the book, one senses Yoxen's feeling of impotence to alter the course of the tide of capitalism which, to give him his due, he recognizes will bring good along with the bad.

What irks him most is that he has seen slip away the hopes kindled by the involvement of trades unions and the public in the early days of genetic engineering. The hopes faded as the business interests moved in and wooed the experts away from righteous paths towards those that were paved with gold; so the argument goes — and with an element of truth.

It is a pity that Yoxen is unwilling to acknowledge the extent to which the interests of profit and social needs seem to go hand in hand, at least in the innovative phase of technology: it is later that so much effort is wasted on redundant products purely for the sake of competition and profit. Nevertheless, in *The Gene Business* Yoxen broaches many of the disturbing questions that tend to be swept aside in the biotechnocrats' rush towards riches. His analysis could have been more penetrating and original, but the book is undoubtedly a provocative read for those who have never stopped to consider such matters at all. □

Peter Newmark is Deputy Editor of Nature.

Biophysics as natural history

R. McNeill Alexander

How Life Learned to Live.

By Helmut Tributsch.

MIT Press: 1983. Pp. 218. \$19.95, £17.95.

PROFESSOR Tributsch had been working at Berkeley on the biophysics of photosynthesis when he came to feel that his subject was losing contact with nature. He had been spending his time operating electronic and optical instruments, reading scientific journals and thinking about abstract models. He felt the need to get out into the sun and see it shining on plants. He wanted to remind himself and others that biophysics is about nature.

So he spent two years travelling in South America — on foot, on horseback, on trucks and on freighters — visiting Amazonia, the Chaco, the Andes, the Galapagos Islands and Tierra del Fuego. He sought out the animals and plants that most of us know only from books and museum specimens, and came back to write his book. It was published in German in 1976 and has now been translated into English.

The book aims to show how animals and plants exploit physical principles. There

are chapters on structural mechanics, locomotion, acoustics, heat, light and electricity. There are passages about beavers' dams, flying seeds, insect song, electric fishes and many other fascinating topics. Professor Tributsch reminds us frequently that he has seen many of the organisms he describes, in their natural habitats. Occasionally he makes good use of his experiences, for example in accounts of the flight and diving of gannets, and of echolocation by oilbirds. More often his reminiscences are mere ornaments in a text that depends more on what he has read than on what he has seen and thought.

He seems to want to amaze us with the marvels of nature. They come out in a rush, example after example, with explanations that are intended to be comprehensible to the layman and never seem to get seriously to grips with the physical principles involved. Physical laws are seldom stated explicitly, and hardly ever quantitatively. Explanations are, therefore, superficial. They are also so condensed that I frequently wondered whether they would

have made any sense to me if I had not already been familiar with the subject. There is too much uncritical acceptance of statements from other books and too many errors such as writing "work" when "power" is meant, or assigning an animal to the wrong phylum. There are however some well-devised diagrams which, properly explained, could have been most revealing.

There are many books designed to teach biologists the most relevant branches of physics, but few that review the biological applications of physics. The ones that I know are books aimed at the mass market, full of slick half-truths and large colour photographs. It is disappointing that Professor Tributsch has not used his expertise in biophysics to break this mould and to write an intellectually more satisfying book. □

R. McNeill Alexander is Professor of Zoology at the University of Leeds. Among other books he is author of *Animal Mechanics* (Blackwell Scientific), which has recently appeared in a second edition.

Sic publishing

Thressa C. Stadtman

Biochemistry of Selenium.

By Raymond J. Shamberger.

Plenum: 1983. Pp. 332. \$42.50, £29.75.

RAYMOND Shamberger's *Biochemistry of Selenium* is the second volume in a series dealing with the biochemistry of the elements, edited by Earl Frieden. In view of the ever-increasing interest in selenium as a micronutrient the book is particularly timely. However it can be recommended only to those readers who wish to own a completely uncritical, annotated bibliography of work involving or related to selenium.

Many of the accounts of experiments are presented in a form garbled to the point of incomprehensibility. The following quotation from p.232, under the heading "Antimutagenicity", is an example:

Sodium selenite and seleno-amino acids interfere with the crossing over in barley (Walker and Ting, 1967). The authors provide cytological evidence for the deformation of the chromatin content of the meiocyte as a result of selenium treatment. Their observations of normal pairing discount the possibility of an effect on the ability of homologs to synapse. The latter two observations lead to the suggestion that selenium may reduce crossingover by a relaxation of the chromosomal protein which may lead to a relaxation stress in the chromosomal fibril. The selenohydryl group is less reactive than the sulfhydryl group (Nickerson *et al.* 1956); hence this and associated differences in bond strengths and distance might introduce alterations in the physicochemical properties of selenosubstituted proteins.

Similarly [*sic*], selenocystine and selenomethionine [*sic*] also interfere with the crossover

distribution along the X chromosome of *Drosophila* [*sic melanogaster* (Ahmed and Walker, 1975). These results were attributed in proteins by the error incorporation of selenocystine residues which supported a mechanism of preexchange DNA breakage induced by stress in an associated protein.

In the area of Shamberger's own expertise, clinical studies are merely enumerated in excessive detail with little or no attempt critically to evaluate the results. Elsewhere in the volume the reader may be informed at the top of the page that a selenium compound in a particular protein is unknown only to be told in the next paragraph of its precise identity. Numerous such examples illustrate the chronological approach which is used throughout the book at the expense of presenting a concise, up-to-date summary of work in this area.

The entire text is full of glaring errors involving the names of chemicals, details of chemical reactions and concentrations of selenium compounds administered, in many instances wrong by several orders of magnitude. For example, on p.254 it is stated that the human patient received 100 mg selenium four times daily for one week and then 25 mg four times daily for the next five days. Had these amounts actually been given the patient should have been dead in 24 hours.

Throughout the book there are incomplete sentences, bad grammar and innumerable spelling errors. Apparently neither the series editor nor the redactory office of the publishing house made any attempt to correct even the most obvious mistakes. □

Thressa C. Stadtman is at the National Heart, Lung, and Blood Institute of the National Institutes of Health, Bethesda, Maryland.

nature
is available in
Microform.



University Microfilms
International

Please send additional information

for _____

Name _____

Institution _____

Street _____

City _____

State _____

Zip _____

300 North Zeeb Road
Dept. P.R.
Ann Arbor, Mi. 48106

Physics

- BOGOLIUBOV, N.N. and SHIRKOV, D.V. *Quantum Fields*. Pp.388. ISBN 0-8053-0983-7. (Addison-Wesley: 1983.) £29.95.
- BREDERLOW, G., FILL, E. and WITTE, K.J. *The High-Power Iodine Laser*. Springer Series in Optical Sciences, Vol. 34. Pp.182. ISBN 3-540-11792-X/0-387-11792-X. (Springer-Verlag: 1983.) DM 87.50, \$35.
- BURT, P.B. *Quantum Mechanics and Nonlinear Waves*. Pp.331. ISBN 3-7186-0072-2. (Harwood Academic: 1981.) \$75.
- CYROT, M. (ed.). *Magnetism of Metals and Alloys*. Pp.608. ISBN 0-444-86518-7. (Elsevier Biomedical/North Holland: 1982.) \$74.50, Dfl. 175.
- DESSLER, A.J. (ed.). *Physics of The Jovian Magnetosphere*. Pp.544. ISBN 0-521-24558-3. (Cambridge University Press: 1983.) £25.
- DRAZIN, P.G. and REID, W.H. *Hydrodynamic Stability*. Pp.527. Hbk ISBN 0-521-22798-4; pbk ISBN 0-521-28980-7. (Cambridge University Press: 1983.) Hbk np; pbk £12.50.
- EINSPRUCH, N.G. (ed.). *VLSI Electronics Microstructure Science*, Vol. 3. Pp.453. ISBN 0-12-234103-1. (Academic: 1982.) \$56.
- EZEKIEL, S. and ARDITTY, H.J. (eds). *Fiber-Optic Rotation Sensors and Related Technologies*. Springer Series in Optical Sciences, Vol.32. *Proceedings of the 1st International Conference held November 1981, Massachusetts*. Pp.440. ISBN 3-540-11791-1/0-387-11791-1. (Springer-Verlag: 1982.) DM 80, \$32.
- KLÉMAN, M. *Points, Lines and Walls: In Liquid Crystals, Magnetic Systems and Various Ordered Media*. Pp.322. ISBN 0-471-10194-X. (Wiley: 1983.) £29.95.
- MACKINNON, E.M. *Scientific Explanation and Atomic Physics*. Pp.450. ISBN 0-226-50053-5. (University of Chicago Press: 1982.) £22.
- MAPLE, M.B. and FISCHER, Ø. (eds). *Superconductivity in Ternary Compounds II: Superconductivity and Magnetism*. Topics in Current Physics, Vol.34. Pp.306. ISBN 3-540-11814-4/0-387-11814-4. (Springer-Verlag: 1982.) DM 79, \$31.60.
- MATSUBARA, T. (ed.). *The Structure and Properties of Matter*. Springer Series in Solid-State Sciences, Vol.28. Pp.446. ISBN 3-540-11098-4/0-387-11098-4. (Springer-Verlag: 1982.) DM 100, \$40.
- MEHRING, M. *Principles of High Resolution NMR in Solids*, 2nd Edn. Pp.342. ISBN 3-540-11852-7/0-387-11852-7. (Springer-Verlag: 1983.) DM 172, \$68.80.
- OHNUKI, Y. and KAMEFUCHI, S. *Quantum Field Theory and Parastatistics*. Pp.489. ISBN 3-540-11643-5/0-387-11643-5. (Springer-Verlag: 1982.) DM 154, \$64.30.
- PIES, W. and WEISS, A. *Landolt-Börnstein: Numerical Data and Functional Relationships in Science and Technology*. Group III: Crystal and Solid State Physics, Vol. 7. Pp.435. ISBN 3-540-11622-2/0-387-11622-2. (Springer-Verlag: 1982.) DM 740, \$296.
- SUNDERMEYER, K. *Constrained Dynamics*. Pp.318. Pbk ISBN 3-540-11947-7/0-387-11947-7. (Springer-Verlag: 1982.) DM33, \$13.20.
- VAN KRAENDONK, J. *Solid Hydrogen: Theory of the Properties of Solid H₂, HD, and D₂*. Pp.306. ISBN 0-306-41080-X. (Plenum: 1983.) Np.

Chemistry

- FORMÁČEK, V. and KUBECZKA, K.-H. *Essential Oils Analysis by Capillary Gas Chromatography and Carbon-13 NMR Spectroscopy*. Pp.373. ISBN 0-471-26218-8. (Wiley: 1982.) £53.50, \$99.95.
- GLASBY, J.S. *Two Part Encyclopaedia of the Terpenoids A-H and I-Z*. Pp.2646. A-H ISBN 0-471-90031-1; I-Z ISBN 0-471-90032-X; 2 part set ISBN 0-471-27986-2. (Wiley: 1982.) £250.
- GOLD, V. and BETHELL, D. (eds). *Advances in Physical Organic Chemistry*, Vol. 19. Pp.428. ISBN 0-12-033519-0. (Academic: 1983.) £37.60, \$69.50.
- GRABMAIER, J. (ed.). *Silicon Chemical Etching*. Crystals, Vol. 8. Pp.226. ISBN 3-540-11862-4/0-387-11862-4. (Springer-Verlag: 1982.) DM 144, \$57.60.

- HARRIMAN, A. and WEST, M.A. (eds). *Photogeneration of Hydrogen*. Pp.210. ISBN 0-12-326380-8. (Academic: 1982.) £12, \$22.50.
- HARTLEY, F.R. and PATAI, S. (eds). *The Chemistry of the Metal-Carbon Bond*, Vol. 1: *The Structure, Preparation, Thermochemistry and Characterization of Organometallic Compounds*. Pp.1071. ISBN 0-471-10058-7. (Wiley: 1982.) £125.
- HARTMANN, H. and WANCZEK, K.-P. *Ion Cyclotron Resonance Spectrometry II. Lecture Notes in Chemistry*, Vol.31. Pp.538. ISBN 3-540-11957-4/0-387-11957-4. (Springer-Verlag: 1982.) Pbk DM 82, \$32.80.
- HINZE, J. (ed.). *Electron-Atom and Electron-Molecule Collisions*. Pp.354. ISBN 0-306-41188-1. (Plenum: 1983.) Np.
- KIRBY, A.J. *The Anomeric Effect and Related Stereoelectronic Effects at Oxygen*. Reactivity and Structure Concepts in Organic Chemistry, Vol.15. Pp.149. ISBN 3-540-11684-2/0-387-11684-2. (Springer-Verlag: 1983.) DM 114, \$45.60.
- Organic Chemistry. Topics in Current Chemistry, Vol. 105. Pp.158. ISBN 3-540-11636-2/0-387-11636-2. (Springer-Verlag: 1982.) DM 86, \$34.40.
- SHAMBERGER, R.J. *Biochemistry of Selenium*. Biochemistry of the Elements, Vol.2. Pp.334. ISBN 0-306-41090-7. (Plenum: 1983.) Np.
- SIMPSON, C.F. (ed.). *Techniques in Liquid Chromatography*. Pp.464. ISBN 0-471-26220-X/0-855-01707-4. (Wiley Heyden: 1982.) £27, \$55.
- Synthetic and Structural Problems. Topics in Current Chemistry, Vol.106. Pp.178. ISBN 3-540-11766-0/0-387-11766-0. (Springer-Verlag: 1982.) DM 86, \$34.40.
- THROWER, P.A. (ed.). *Chemistry and Physics of Carbon*, Vol. 18. Pp. 208. ISBN 0-8247-1740-6. (Dekker: 1982.) SwFr. 135.
- WARREN, S. *Organic Synthesis: The Disconnection Approach*. Pp.391. Hbk ISBN 0-471-10160-5; pbk ISBN 0-471-10161-3. (Wiley: 1983.) Hbk np; pbk £7.95, \$15.95.
- WEBER, W.P. *Silicon Reagents for Organic Synthesis*. Reactivity and Structure Concepts in Organic Chemistry, Vol. 14. Pp.430. ISBN 3-540-11675-3/0-387-11675-3. (Springer-Verlag: 1983.) DM 224, \$89.60.
- BORGESE, E.M. and GINSBURG, N. (eds). *Ocean Yearbook 3*. Pp.581. ISBN 0-226-06604-5. (University of Chicago Press: 1982.) £39.20.
- LARSEN, G. and CHILLINGAR, G.V. (eds). *Diagenesis in Sediments and Sedimentary Rocks*, 2. *Developments in Sedimentology 25B*. Pp.572. ISBN 0-444-42013-4. (Elsevier Biomedical/North Holland: 1983.) £117, Dfl. 275.
- MELCHIOR, P. *The Tides of the Planet Earth*, 2nd Edn. Pp.641. ISBN 0-08-026248-1. (Pergamon: 1983.) £45, \$90.
- SCHLANGER, S.O. and CITA, M.B. (eds). *Nature and Origin of Cretaceous Carbon-rich Facies*. Pp. 240. ISBN 0-12-624950-4. (Academic: 1982.) \$21, \$39.
- VAN ROSE, S. *Earthquakes*. Pp.36. Pbk ISBN 0-11-884066-5. (HMSO: 1983.) 90p.

Earth Sciences

Environmental Sciences

- HILL, R.B. and TERZIAN, J.A. (eds). *Environmental Pathology: An Evolving Field*. Pp.366. ISBN 0-8451-0221-4. (Alan R. Liss: 1982.) £46.
- HOLZNER, W., WERGER, M.J.A. and IKUSIMA, I. (eds). *Man's Impact on Vegetation*. Geobotany, Vol. 5. Pp.370. ISBN 90-6193-685-3. (Dr. W. Junk: 1983.) DG 225, \$98.
- KOLBER, A.R. *et al.* (eds). *In Vitro Toxicity Testing of Environmental Agents: Current and Future Possibilities*, Part A: Survey of Test Systems. Nato Conference Series, Vol. 5. Pp.553. ISBN 0-306-41123-7. (Plenum: 1983.) Np.
- KOLBER, A.R. *et al.* (eds). *In Vitro Toxicity Testing of Environmental Agents: Current and Future Possibilities*, Part B: Development of Risk Assessment Guidelines. Nato Conference Series, Vol. 5B. Pp.545. ISBN 0-306-41124-5. (Plenum: 1983.) Np.

Technology

- ANDERSON, J.R. and BOUDART, M. (eds). *Catalysis: Science and Technology*, Vol. 3. Pp.289. ISBN 3-540-11634-6/0-387-11634-6. (Springer-Verlag: 1982.) DM 128, \$51.20.
- CARRAHER, C.E. Jr. and SPERLING, L.H. (eds). *Polymer Applications of Renewable-Resource Materials*. Polymer Science and Technology, Vol. 17. Pp.469. ISBN 0-306-41033-8. (Plenum: 1983.) Np.
- HAPPEY, F. (ed.). *Contemporary Textile Engineering*. Pp.454. ISBN 0-12-323750-5. (Academic: 1982.) £39.80, \$74.
- KUTSCHER, C.F. *et al.* *Design Approaches for Solar Industrial Process Heat Systems: Nontracking and Line-Focus Collector Technologies*. Pp.423. Pbk ISBN 0-89553-113-5. (American Solar Energy Society: 1982.) \$54.
- ROPER, C.C. and SILVA, J. (eds). *Science and Technology in Latin America*. Pp.384. ISBN 0-582-90057-3. (Longman: 1983.) £39.
- SCULLY, J.C. (ed.). *Corrosion: Aqueous Processes and Passive Films*. Treatise on Materials Science and Technology, Vol. 23. Pp.480. ISBN 0-12-633670-9. (Academic: 1983.) £38.80, \$72.

Biological Sciences

- ASHTON, H. and STEPNEY, R. *Smoking: Psychology and Pharmacology*. Pp.222. Hbk ISBN 0-422-77700-5; pbk ISBN 0-422-77710-2. (Tavistock: 1983.) Hbk np; pbk £3.95.
- BOURLIÈRE, F. (ed.). *Tropical Savannas. Ecosystems of the World*, Vol. 13. Pp.730. ISBN 0-444-42035-5. (Elsevier Biomedical/North Holland: 1983.) \$197.75, Dfl. 465.
- BURGHARDT, G.M. and RAND, A.S. (eds). *Iguanas of the World: Their Behaviour, Ecology, and Conservation*. Pp.472. ISBN 0-8155-0917-0. (Noyes: 1982.) \$55.
- BUSCH, H. and ROTHBLUM, L. (eds). *The Cell Nucleus*, Vol. 12: rDNA, Part C. Pp.248. ISBN 0-12-147612-X. (Academic: 1982.) \$42.
- CAMPBELL, P.N. and MARSHALL, R.D. *Essays in Biochemistry*, Vol. 18. Pp.126. Pbk ISBN 0-12-158118-7. (Academic: 1982.) £8, \$12.50.
- COTLIER, E., MAUMENEE, I.H. and BERMAN, E.R. (eds). *Genetic Eye Diseases: Retinitis Pigmentosa and Other Inherited Eye Disorders*. *Proceedings of the International Symposium held September 1981, Jerusalem. Birth Defects: Original Article Series*, Vol. 18, No. 6. Pp.772. ISBN 0-8451-1050-0. (Alan R. Liss: 1982.) £58.
- DALY, J.W. *et al.* (eds). *Physiology and Pharmacology of Adenosine Derivatives*. Pp.313. ISBN 0-89004-833-9. (Raven: 1982.) \$71.92.
- DAUSSANT, J., MOSSÉ, J. and VAUGHAN, J. (eds). *Seed Proteins. Annual Proceedings of the Phytochemical Society of Europe*, No. 20. Pp. 350. ISBN 0-12-204380-4. (Academic: 1983.) £28, \$50.
- FEIN, A. and SZUTS, E.Z. *Photoreceptors: Their Role in Vision*. Pp.212. Hbk ISBN 0-521-24433-1; pbk ISBN 0-521-28684-0. (Cambridge University Press: 1983.) Hbk £17.50; pbk £7.95.
- GOVINDJEE, A. (ed.). *Photosynthesis*, Vol.1: *Energy Conversion by Plants and Bacteria*. Pp.799. ISBN 0-12-294301-5. (Academic: 1982.) \$79.
- GRACE, J. *Plant-Atmosphere Relationships*. Pp.92. Pbk ISBN 0-412-23180-8. (Chapman & Hall: 1983.) £2.95.
- HENCH, L.L. and ETHRIDGE, E.C. *Biomaterials: An Interfacial Approach*. Pp.385. ISBN 0-12-340280-8. (Academic: 1982.) \$42.
- HURLBERT, S.H. and VILLALOBOS-FIGUEROA, A. (eds). *Aquatic Biota of Mexico, Central America and the West Indies*. Pp.529. No ISBN. (San Diego State University: 1982.) \$20.
- HURVICH, L.M. *Color Vision*. Pp.328. Hbk ISBN 0-87893-336-0; pbk ISBN 0-87893-337-9. (Sinauer: 1981.) £16.50.
- KENNEDY, J.F. and WHITE, C.A. *Bioactive Carbohydrates: In Chemistry, Biochemistry and Biology*. Pp.331. Hbk ISBN 0-85312-201-6; pbk ISBN 0-85312-467-1/0-470-27527-8. (Ellis Horwood/Halsted: 1983.) Hbk £32.50, \$65.35; pbk np.
- KLAHR, S. (ed.). *The Kidney and Body Fluids in Health and Disease*. Pp.593. ISBN 0-306-41062-1. (Plenum: 1983.) Np.

KUO, P.T. CONN, H.L. Jr. and DeFELICE, E.A. (eds). *Health and Obesity*. Pp.200. ISBN 0-89004-809-6. (Raven: 1982.) \$36.58.

MARSHALL, K.C. (ed.). *Advances in Microbial Ecology*, Vol. 6. Pp. 240. ISBN 0-306-41064-8. (Plenum: 1982.) \$39.50.

MEUDT, W.J. (ed.). *Strategies of Plant Reproduction*, Vol. 6. Invited Papers Presented at a Symposium held May 1981, at the Beltsville Agriculture Research Center, Beltsville. Pp. 400. ISBN 0-86598-054-3/0-246-11946-2. (Granada/Allanheld, Osmun & Co: 1983.) £25.

MIMS, C.A. *The Pathogenesis of Infectious Disease*, 2nd Edn. Pp. 297. Hbk ISBN 0-12-498254-9; pbk ISBN 0-12-498455-X (Academic: 1982.) Hbk np; pbk £5.60, \$10.50.

MISAGHI, I.J. *Physiology and Biochemistry of Plant-Pathogen Interactions*. Pp. 287. ISBN 0-306-41059-1. (Plenum: 1982.) \$32.50.

MIZUNO, D. *et al.* (eds). *Self-Defense Mechanisms: Role of Macrophages*. Pp. 343. ISBN 0-444-80460-9. (Elsevier Biomedical/North Holland: 1982.) \$83, Dfl. 195.

MORIN, R.B. and GORMAN, M. (eds). *Chemistry and Biology of β -Lactam Antibiotics*, Vol. 1: Penicillins and Cephalosporins. Pp. 553. ISBN 0-12-506301-6. (Academic: 1982.) \$73.

MULLAN, F. *Vital Signs: A Young Doctor's Struggle with Cancer*. Pp. 203. No ISBN. (Farrar, Straus & Giroux: 1983.) \$12.50.

MUNSON, P.L. *et al.* (eds). *Vitamins and Hormones: Advances in Research and Applications*. Pp. 431. ISBN 0-12-709839-9. (Academic: 1982.) \$56.

NEFF, W.D. (ed.). *Contributions to Sensory Physiology*, Vol. 6. Pp. 209. ISBN 0-12-151806-X. (Academic: 1982.) \$39.50.

Psychology

BARBER, J. and ADRIAN, C. (ed.). *Psychological Approaches to the Management of Pain*. Pp.230. ISBN 0-87630-303-3. (Brunner/Mazel: 1982.) \$20.

BLANKSTEIN, K.R. and POLIVY, J. (eds). *Self-Control and Self-Modification of Emotional Behaviour*. *Advances in the Study of Communication and Affect*, Vol. 7. Pp.204. ISBN 0-306-40945-3. (Plenum: 1982.) Np.

GUGGENHEIM, F.G. and NADELSON, C. (eds). *Major Psychiatric Disorders: Overviews and Selected Readings*. Pp.388. Pbk ISBN 0-444-00663-X. (Elsevier Biomedical/North Holland: 1982.) \$19.95, Dfl. 65.

HESLIN, R. and PATTERSON, M.L. *Nonverbal Behaviour and Social Psychology*. Pp.183. ISBN 0-306-40952-6. (Plenum: 1982.) \$24.50.

LAHEY, B.B. and KAZIN, A.E. (eds). *Advances in Clinical Child Psychology*, Vol. 5. Pp.375. ISBN 0-306-41043-5. (Plenum: 1982.) \$35.

LOWENTHAL, F., VANDAMME, F. and CORDIER, J. (eds). *Language and Language Acquisition*. Pp.379. ISBN 0-306-41128-8. (Plenum: 1982.) \$42.50.

McGUIGAN, F.J. *Calm Down: A Guide to Stress and Tension Control*. Pp.230. Hbk ISBN 0-13-112839-6; pbk ISBN 0-13-112821-3. (Prentice-Hall: 1982.) Hbk np; pbk £4.75.

MILLON, T., GREEN, C. and MEAGHER, R. (eds). *Handbook of Clinical Health Psychology*. Pp.608. ISBN 0-306-40932-1. (Plenum: 1982.) \$50.

History of Science

CHANDLER, B. and MAGNUS, W. *The History of Combinatorial Group Theory: A Case Study in the History of Ideas*. *Studies in the History of Mathematics and Physical Sciences*, Vol. 9. Pp.234. ISBN 0-387-90749-1/3-540-90749-1. (Springer-Verlag: 1982.) DM 128, \$51.20.

FEDERICO, P.J. *Descartes on Polyhedra: A Study of the De Solidorum Elementis*. *Sources in the History of Mathematics and Physical Sciences*, Vol. 4. Pp.145. ISBN 0-387-90760-2/3-540-90760-2. (Springer-Verlag: 1982.) DM 98, \$39.20.

MOORE, G.H. *Zermelo's Axiom of Choice: Its Origins, Development, and Influence*. *Studies in the History of Mathematics and Physical Sciences*, Vol. 8. Pp.410. ISBN 0-387-90670-3/3-540-90670-3. (Springer-Verlag: 1982.) DM 108, \$43.20.

NAGEL, E. *Teleology Revisited and Other Essays in the Philosophy and History of Science*. Pp.352. Hbk ISBN 0-231-04504-2; pbk ISBN 0-231-04505-0. (Columbia University Press: 1982.) Hbk np; pbk \$17.

SESIANO, J. *Books IV to VII of Diophantus' Arithmetica in the Arabic Translation Attributed to Qustā ibn Lūqā*. *Sources in the History of Mathematics and Physical Sciences*, Vol. 3. Pp.502. ISBN 0-387-90690-8/3-540-90690-8. (Springer-Verlag: 1982.) DM 194, \$77.60.

Anthropology

BLANTON, R.E. *et al.* *Monte Alban's Hinterland, Part 1: The Prehispanic Settlement Patterns of the Central and Southern Parts of the Valley of Oaxaca, Mexico*, Vol. 7. *Memories of the Museum of Anthropology*, University of Michigan, No.15. Pp.506. Pbk ISBN 0-932206-91-3. (Regents of the University of Michigan: 1982.) Np.

MOLNAR, S. *Human Variation: Races, Types, and Ethnic Groups*, 2nd Edn. Pp.253. Pbk ISBN 0-13-447664-6. (Prentice-Hall: 1983.) £13.55.

RENFREW, C., ROWLANDS, M.J. and SEGRAVES, B.A. (eds). *Theory and Explanation in Archaeology: The Southampton Conference*. Pp.480. ISBN 0-12-586960-6. (Academic: 1982.) \$49.50.

General

AIDA, S. (ed.). *The Humane Use of Human Ideas: The Discoveries Project and Eco-Technology*. Pp.311. Hbk ISBN 0-08-027941-4; pbk ISBN 0-08-027942-2. (Pergamon: 1983.) Hbk £20, \$40; pbk £9.95, \$19.95.

BEITZ, D.C. and HANSEN, R.G. (eds). *Animal Products in Human Nutrition*. Pp.545. ISBN 0-12-086380-4. (Academic: 1982.) \$62.

BERRY, A. *The Super-Intelligent Machine: An Electronic Odyssey*. Pp.182. ISBN 0-224-01967-8. (Jonathan Cape: 1983.) £7.95.

CHOW, V.T. (ed.). *Advances in Hydrospace*, Vol. 13. Pp.393. ISBN 0-12-021813-5. (Academic: 1982.) \$53.

Expensive Medical Techniques: Report of a Working Party. Pp.59. Hbk ISBN 0-906503-11-6; pbk ISBN 0-906503-12-4. (Council for Science and Society: 1983.) Hbk £5; pbk £2.50.

FLOOD, M. *Solar Prospects: The Potential for Renewable Energy*. Pp.190. Pbk ISBN 0-7045-0473-1. (Wildwood House Ltd: 1983.) £6.95.

FROST, H.J. and ASHBY, M.F. *Deformation-Mechanism Maps: The Plasticity and Creep of Metals and Ceramics*. Pp.250. Hbk ISBN 0-08-029338-7; pbk ISBN 0-08-029337-9. (Pergamon: 1982.) Hbk £30, \$60; pbk £15, \$30.

FULDER, S. *An End to Ageing? Traditional and Modern Ways of Extending Healthy Life*. Pp.112. Pbk ISBN 0-7225-0769-0. (Thorsons, Northampton: 1983.) £3.50.

GEYER, R.F. and VAN DER ZOUWEN, J. (eds). *Dependence and Inequality: A Systems Approach to the Problems of Mexico and Other Developing Countries*. Pp.326. ISBN 0-08-027952-X. (Pergamon: 1982.) £17.50, \$35.

GOFF, P.E. *Nature, Children and You*. Pp.144. Hbk ISBN 0-8214-0607-8; pbk ISBN 0-8214-0679-5. (Ohio University Press: 1982.) £10.50.

GRANT, M., PLANT, M. and WILLIAMS, A. (eds). *Economics and Alcohol: Consumption and Controls*. Pp.302. ISBN 0-7099-1132-7/0-89876-089-5. (Croom Helm/Gardner Press: 1983.) £16.95.

HAGAN, P.J. *Blood: Gift or Merchandise: Towards an International Blood Policy*. Pp.231. ISBN 0-8451-0219-2. (Alan R. Liss: 1982.) £17.

HILGARTNER, S., BELL, R.C. and O'CONNOR, R. *Nukespeak: Nuclear Language, Visions, and Mindset*. Pp.282. ISBN 0-87156-307-X. (Sierra Club, California: 1982.) \$14.95.

HOLLIS, M. and LUKES, S. (eds). *Rationality and Relativism*. Pp.312. Hbk ISBN 0-631-12773-9; pbk ISBN 0-631-13126-4. (Basil Blackwell Publisher: 1982.) Hbk £16; pbk £6.50.

HUBBARD, R., HENIFIN, M.S. and FRIED, B. (eds). *Biological Woman — The Convenient Myth: A Collection of Feminist Essays and a Comprehensive Bibliography*. Pp.376. Hbk ISBN 0-87073-702-3; pbk ISBN 0-87073-703-1. (Schenkman: 1982.) Hbk \$18.95; pbk \$11.95.

HUBBLE, E. *The Realm of the Nebulae*. Pp.207. Hbk ISBN 0-300-02499-1; pbk ISBN 0-300-02500-9. (Yale University Press: 1983.) Hbk £25; pbk np.

HUISINGH, D. and BAILEY, V. (eds). *Making Pollution Prevention Pay: Ecology with Economy as Policy*. Pp.156. ISBN 0-08-029417-0. (Pergamon: 1982.) £12.50, \$25.

HUNT, M. *The Universe Within: A New Science Explores the Human Mind*. Pp.415. ISBN 0-7108-0437-7. (Harvester Press: 1982.) £12.95.

JEFFERIES, R. *The Life of the Fields*. Pp.224. Pbk ISBN 0-19-281358-7. (Oxford University Press: 1983.) £2.95.

LANGFORD, T.E. *Electricity Generation and the Ecology of Natural Waters*. Pp.342. Pbk ISBN 0-85323-334-9. (Liverpool University Press: 1983.) £14.50.

LEACOCK, E. and LEE, R. (eds). *Politics and History in Band Societies*. Pp.500. UK Hbk ISBN 0-521-24063-8; pbk ISBN 0-521-28412-0; France ISBN 2-7351-0019-7. (Cambridge University Press: 1982.) Hbk £29.50; pbk £9.95.

LÉVI-STRAUSS, C. *The Way of the Masks*. Pp.249. ISBN 0-224-02081-1. (Jonathan Cape: 1983.) £15.

MACRAE, R. (ed.). *HPLC in Food Analysis*. Pp.330. ISBN 0-12-464780-4. (Academic: 1982.) £28, \$50.

McPHEE, J. *Basin and Range*. Pp.216. ISBN 0-374-10914-1. (Farrar, Strauss & Giroux, New York: 1983.) £6.

PIGGOTT, S. *Scotland Before History*. Pp.195. ISBN 0-85224-348-0. (Edinburgh University Press: 1982.) Hbk £11.50; pbk £5.75.

PYATT, E. *The National Physical Laboratory: A History*. Pp.270. ISBN 0-85274-387-4. (Adam Hilger: 1983.) £24.

READER, J. *Kilimanjaro*. Pp. 84. ISBN 0-241-10683-4. (Elm Tree, London: 1982.) £12.95.

ROBINSON, D.W. and MOLLAN, R.C. (eds). *Energy Management and Agriculture. Proceedings of the 1st International Summer School*. Pp. 441. ISBN 0-86027-010-6. (Royal Dublin Society: 1982.) IR£25.

ROTHSCHILD, J. (ed.). *Women, Technology and Innovation*. Pp. 289-382. ISBN 0-08-028943-6. (Pergamon: 1982.) £9.50, \$19.

RUBNER, M. *A Nutrition Foundations' Reprint of The Laws of Energy Consumption in Nutrition*. Pp.371. ISBN 0-12-602150-3. (Academic: 1982.) \$42.50.

SAFE DRINKING WATER COMMITTEE. *Drinking Water and Health*, Vol. 4. Pp. 299. Pbk ISBN 0-309-03198-2. (National Academy Press: 1982.) \$14.50, \$19.25.

Sampling and Analysis for the Mineral Industry. Papers Presented at a Symposium held November 1982, London. Pp. 119. Pbk ISBN 0-900488-64-6. (Institution of Mining and Metallurgy: 1982.) £15.

Scientific Basis of Water-Resource Management. Pp. 127. Pbk ISBN 0-309-03244-X. (National Academy Press: 1982.) £11, \$14.75.

SHORTER, J. *Correlation Analysis of Organic Reactivity. With Particular Reference to Multiple Regression*. Pp. 235. ISBN 0-471-10479-5. (Wiley: 1982.) £18.50, \$39.95.

SMART, I. (ed.). *World Nuclear Energy: Toward a Bargain of Confidence*. Pp. 394. ISBN 0-8018-2652-7. (The Johns Hopkins University Press: 1982.) \$30.

SOLOMON, K. and LEVY, N.B. (eds). *Men in Transition: Theory and Therapy*. Pp.503. ISBN 0-306-40976-3. (Plenum: 1982.) \$39.50.

STIBIC, V. *Tools of the Mind: Techniques and Methods for Intellectual Work*. Pp. 300. 0-444-86444-X. (Elsevier Biomedical/North Holland: 1982.) \$35, Dfl. 85.

The Scientific Importance of Submillimetre Observations. Proceedings of a Workshop held May 1982, The Netherlands. Pp. 232. No ISBN. ISSN 0379-6566. (European Space Agency: 1982.) 140 FF.

The World of Philosophy, Vol. 3. Pp. 508. ISBN 3-8055-3419-1. (Karger: 1982.) SwFr. 297, DM 356, \$178.

TRENARD, Y. *Making Wood Speak: An Introduction to Dendrochronology*. *Forestry Abstracts. Review Article*, Vol. 43. No. 12. No ISBN. (Commonwealth Forestry Bureau: 1982.) Pbk £5.50.

TROWELL, H.C., DAVIES, J.N.P. and DEAN, R.F.A. *A Nutrition Foundations' Reprint of Kwashiorkor*. Pp.308. ISBN 0-12-701150-1. (Academic: 1982.) \$35.

On-line digital imaging

Ed Boyes looks at digital image processing — the technique which gleans quantitative data from pictures.

DIGITAL techniques are increasingly being used to enhance, standardize and sometimes replace conventional photographic recordings. This growing interest in digital analysis arises largely from the fact that image processing puts a quantitative value on an otherwise qualitative picture, making it possible to compare two photographs accurately, or to glean more detailed information from a single image.

The essential difference between conventional photographs and digitally-processed images is that in the latter the image is converted to a digital form and stored as numbers in a large semiconductor image memory, or "framestore" attached to a computer, rather than being exposed directly onto a piece of film or an instant film pack. With typical commercially available equipment this means that images of at least 512×512 picture points, or pixels, can be quantified and stored directly from virtually any input, such as a TV camera or an ultra-sound scanner. The picture can be analysed with 8-bit precision, by splitting the image into 2^8 (or 256) shades of grey or colour before it is stored, with a chance of "doctoring" the image both immediately after input and again before output. In this way the image can be changed before an actual picture is produced, so that areas of interest can be enhanced and the background subdued, or colours can be added to increase contrast. Image processing also provides the opportunity to present other physical values in a visual form.

Advantages

Image enhancement procedures are designed to produce the best possible image from the input, and to assist the analysis and communication of results. In doing so, digital techniques have three major advantages over conventional techniques. First, once the data have been in the computer, the images are available for processing in any order with random access to the memory, unlike conventional films which have to be analysed in sequence. Second, image sequences can be displayed on the video screen and frozen electronically at any time for closer inspection. Third, the original data is kept intact and can be repeatedly analysed without distortion. Furthermore, any analytical procedures used on the data can be explicitly defined and recorded.

Enhancement techniques

At the most simple, but nevertheless essential, level of data processing, the brightness and contrast of the image are enhanced so that features of special interest can be picked out more easily. More complicated imaging needs more advanced techniques. Many scientific signals are either very infrequent or very weak, coming complete with heavy background noise. To break these down requires long integration times, during which the image is built up in the memory store and controlled from the display on a video screen. Signal to noise ratios can be improved by averaging successive frames from a TV camera with a controlled output or by summing the data from adjacent pixels, which brings out the similarities in the data and hence the differences between the data and the background. When the noise level of the input is lower, a modified processing of adjacent pixels, for example in a 3×3 array, can be used to sharpen edge-detail by enhancing the differences rather than the similarities between the groups. Similarly, information about intensity may be made more meaningful by producing a non-linear output. An example of this would be in an experiment where the temperature was being increased from 0 to $1,000^\circ\text{C}$, but in which it was important to see the point at which 600°C was reached. This changeover would be made more obvious by representing, for example, all temperatures below 599°C by the colour green and all those above 600°C by the colour red.

Clearly, digital image processing is a useful analytical technique. It therefore seems curious that it has only become widely used in the last few years, even more so when you consider that analog electronic analysis has been in use far longer, despite the fact that it can only analyse serially and the information that it gives is much less useful. The problem with developing digital techniques has been that until recently, digital image processing has required a substantial mainframe computer, and, even then, there were only facilities for on-line analysis and control in very specialized applications. Now, however, laboratory-scale systems are available with typical image memories of 512×512 or $1,024 \times 1,024$ pixels. Each picture point from the input is mapped to an equivalent position on the computer

video screen. If the input frequency is slow, the image processing can be carried out by computer alone. This gives a highly flexible system, but severely limits the rate at which data can be handled. To reach the 15 MHz required by a raster scan demands that some flexibility be sacrificed by the introduction of a piece of dedicated hardware, such as the recursive video processor, or RVP.

The RVP is essentially capable of doing three things — it can take the difference between adjacent frames, and it can sum or average incoming signals to improve the signal to noise ratio. Averaging the signal involves taking data from the store and comparing it with the incoming data, so that, for example, one eighth of the output comes from the input data, and seven eighths from the running average. In addition to controlling noise, it is important to avoid distortion of the image during processing. Any digital processing system is capable of generating a picture without an input, and so it is essential to make sure that interference is kept to a minimum.

It seems that in the future, digital imaging techniques will split into three main categories with cost being their main divider. First, there will be the simple units to store or compare single images in digital form with a TV output display. These are now becoming available as single computer cards, similar to those upon which the latest high-resolution colour VDUs are based. Second, when any significant level of processing becomes necessary, some kind of computer will have to be involved, and this in turn will require extensive software, suggesting that some kind of package system will be developed. At the third level will come the highly sophisticated, fully integrated systems with powerful processors attached to multiple framestores and extensive processing, analysis and storage facilities.

Luckily, it is emerging that many image processing methods require very similar hardware and software. It appears that all that is needed to specialize the systems is a tailor-made sensor coupled to personalized software. □

Ed Boyes is from the Department of Metallurgy and Science of Materials of the University of Oxford, UK.

Digital image analysis systems

● An image analysis system, the Intellect Retina system, has been launched by Micro Consultants for use in situations where high resolution images are required at high speeds. Potential applications include inspection on industrial production lines, and the monitoring of laboratory experiments where low cycle changes are marked by electromagnetic, auditory or thermal outputs. The system is available in two forms — a black box unit and a development system version — and these can be used either for fully automatic operation or with an operator. Unlike other image analysis systems, the Intellect Retina has a 512 line $\times$ 512 pixel resolution and operates over 256 grey levels. It consists of Micro Consultants' real time analyser, together with their Intellect 100 framestore, recursive video processor and 68000-based computer. Software for the system is blown into prom, resident within the black box, allowing the system to operate independently. The development systems version of the Intellect Retina includes the black box system, having its computer in a DEC VT103 terminal, either in the form of an LSI-11 or a Q-bus-compatible 68000. A DSD 880 Winchester disk peripheral is also included, providing facilities for program loading, development and storage. Program development can be carried out in Fortran or Pascal. When the software has been finalized, proms can be blown accordingly for the 68000 in the black box unit for production line use.

Circle No. 100 on Reader Service Card

● Gould has produced a high speed digital disk which allows image processing systems to view high-resolution displays at the real time data transfer rates of most video devices. The Gould DeAnza real-time digital disk uses 200 Mbytes of virtual image storage to record and play-back 512 $\times$ 512 $\times$ 10 bit imagery at 30 frames per second. The real time digital disk allows the recording and play-back of up to 800 inter-related images. The system has applications in fluoroscopy, angiography, seismic data analysis, structure analysis, animation, and modelling. The real-time digital disk features 300 Mbytes of modified Winchester disk, two Hex board set controllers, one Hex board interface to Gould DeAnza IP 8500, IP 8400, and IP 7400 image processors. The disk can be formatted dynamically as 800 virtual 256-byte cylinders, 1,600 virtual, 128-kbyte cylinders, or 3,200 virtual, 64-kbyte cylinders.

Circle No. 101 on Reader Service Card

These notes are based on information provided by the manufacturers. For further details circle the appropriate numbers on the Reader Service Card bound inside the journal.

● Scanning electron microscope images can now be analysed automatically using Link Systems new Digiscan package for feature detection and classification. The Link image analysis system has been designed primarily to integrate chemical classification based on energy-dispersive X ray analysis with the capability to detect, count and size features of interest within the field of view. Efficiency can be enhanced substantially by automatic pre-selection of features such as particles and inclusions which are likely to be of analytical importance. Electronic hardware includes a digital scan generator with a standard capability of 1,024 $\times$ 1,024 and optional 4,096 $\times$ 4,096 grid point resolution, a precision signal digitizer and a high-speed window comparator. After all features have been counted and sized, the beam is scanned only over those features whose dimensions fulfil specified conditions and an X ray analysis is performed for each of these features. The analysis program generates a disk file containing a summary of pertinent information for each detected feature. This disk file is interrogated, when all acquisition operations have been completed, in order to sort features into types based on shape, size and chemical class defined by specified concentration limits, so that relevant statistical information can be displayed. The Digiscan package is designed for routine counting.

Circle No. 102 on Reader Service Card

● A digital imaging workstation for the image processing industry is now being produced by International Imaging Systems. This system 575 is provided with an image processing software package, a model 75 digital image processor and an imbedded computer. The system has an 80 Mbyte disk drive, a 1,600 bits per inch (b.p.i.) magnetic tape system, and a VT-100 terminal.

Circle No. 103 on Reader Service Card.



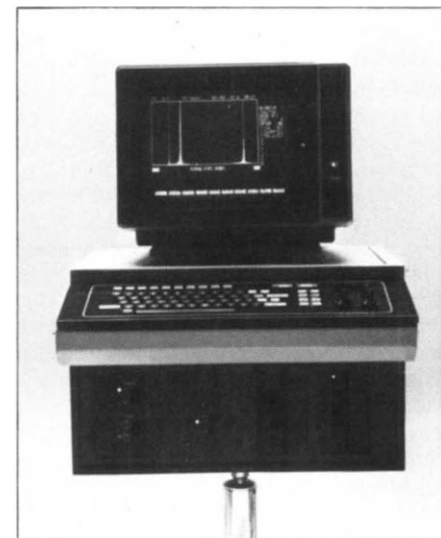
Link System's Digiscan package

● A new microanalytical technique that enables a scanning electron microscope to analyse the composition of compounds containing elements that cannot be detected by energy dispersive X ray analysis (EDS) or wet chemical methods has been introduced by International Scientific Instruments. The new system, called Card (compositional analysis using a Robinson detector), consists of a microprocessor-based analyser and an atomic-number-sensitive, retractable Robinson backscattered electron (RBSE) detector to classify a specimen or phase according to its average atomic number. It then uses that information to identify the material by its compound formula. This method can be used to analyse compounds containing light elements such as hydrogen and fluorine since even slight differences in average atomic number can be detected and used to determine composition. Card comes fully equipped with all the hardware and electronics required to install it on any scanning electron microscope. The workstation includes a dual floppy disk drive, a keyboard and a colour display terminal. Software is also provided with a range of analysis programs for use in identifying compounds and performing calculations. Card allows rapid spot analysis and full frame analysis of all compounds within a field of view.

Circle No. 104 on Reader Service Card.

● The Scopix 300R imager from Agfa-Gevaert is designed to accept a video input from which hard photographic copies are produced. The Scopix 300R has an integrated flat-face monitor and its own twin lens optical system, which enables photographs to be produced. The Scopix 300R will accept two independent video signals and will be linked to an RS-232 computer interface. The imager uses daylight-loading rolls of Agfa-Gevaert film or paper. The use of long rolls allows continuous operation, remote control and computer linking.

Circle No. 105 on Reader Service Card



The Card System from ISI

Cameras

● Two appliances are now being produced by Polaroid for use in developing micrographs. The first, the CU-5 hard-copy camera, takes records of oscilloscope, video monitor, computer terminal and scanning electron microscope traces. Standard focusing hoods of $3\frac{1}{8} \times 3\frac{7}{8}$ and $6\frac{1}{2} \times 8\frac{3}{4}$ inches are available, although larger hoods can be commissioned to order. Two models of the CU-5 — the $3\frac{1}{4} \times 4\frac{1}{4}$ inches and the 4×5 inches — can make both colour and black and white prints either with or without a negative. The second appliance from Polaroid is an adapter for the SX-70 camera which allows any folding Polaroid SX-70 camera with a tripod mount to be fitted to most optical microscopes.

Circle No. 106 on Reader Service Card

● The Eikonixscan 78/99 is a digital camera with modular options. It digitizes object material, as well as transparent and opaque media, on formats ranging from fractions of an inch to several feet. The camera can be used in conjunction with a standard microscope, and can be run through its microprocessor-based control either by a host computer, or it can be operated alone. Options available include illumination systems, camera stands, computer interfaces and signal processing boards.

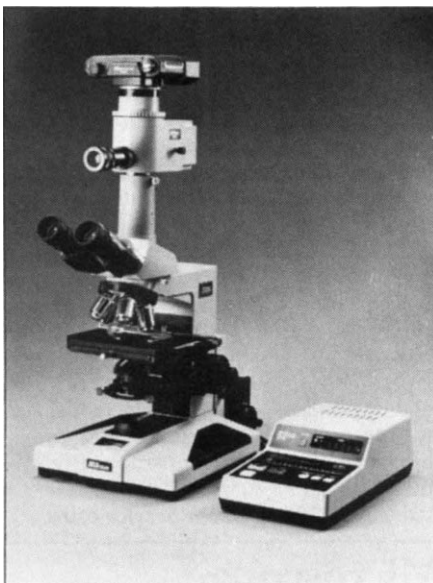
Circle No. 107 on Reader Service Card.

● Wild Heerbrugg has developed four Wild MPS photomicrography systems which, when used in conjunction with a Leitz camera, will automatically give photomicrographs and macrographs. Settings for the camera factor (for formats from 36×24 mm to 4×5 inches), exposure metering (for films from 5 to 43 DIN), exposure time and film transport are carried out automatically. Exposure time is displayed on the control unit and can be overridden by a manual control.

Circle No. 108 on Reader Service Card.

● Ultra-Violet Products (UVP) is launching a camera system for researchers using transilluminators and UV viewing cabinets. UVP now offers two complete camera systems for UV research based on the Shackman 7000 and Polaroid CU-5 cameras. The system requires no setting: the gel is placed on the filter glass, and the camera and hood assembly put on top. Custom-built hoods and corrector lenses are available for most makes of camera. Each system can be bought complete or in kit form for use with existing transilluminators or UV cabinets. When a UVP CC-60 Chromato-Vue cabinet is being used — alone or in conjunction with a transilluminator — either camera system can be attached to the cabinet's viewport.

Circle No. 109 on Reader Service Card



Nikon's new photomicrographic series

● A new photomicrographic series has been launched by Nikon. This new microflex FX series is for use in conjunction with a Nikon microscope. The series includes the fully automatic microflex UFX with a built-in computerized control system; the microflex HFX and the microflex AFX, both of which are more basic automatic microflexes giving 30% integrated measurement compared with the 1% offered by the UFX; and the microflex PFX, which is a standard microflex that fits the vertical phototube. A number of accessories are also available, including a motorized dark box (FX-35A) which is compatible with the UFX and HFX, a plain dark box (FX-35) which is compatible with the AFX and PFX for 35 mm film, a large format adapter (4X) which can be mounted onto the FX series microflexes, an ocular finder and a focusing magnifier.

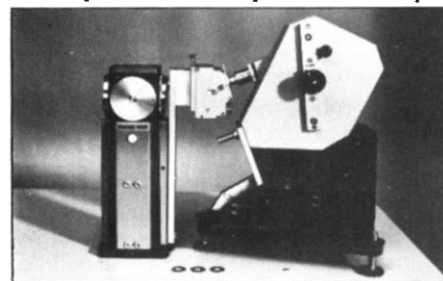
Circle No. 110 on Reader Service Card.

● Olympus now produce a series of photomicrographic equipment, from the manually-operated PM-6 35 mm camera to the fully automatic PM-10AD system camera. The PM-10AD system has a built-in microcomputer which will calculate the correct exposure time based on information fed in by the operator on the format size, film sensitivity and reciprocity failure characteristics of the specimen. A time lapse control unit, the PM-IV, can also be attached. The second in the series is the PM-10M, a manual photomicrography system with an attachable data imprinter. Also available is the PM-6, an attachment camera which can be fitted to any microscope with a 25 mm external phototube. The last in the series is the PMT-35 photomicrography system, which is a macro-adaptor designed for the Olympus SLR camera system. It has three macro-lenses — 80 mm f4; 38 mm f3.5; and 20 mm f3.5 — which make it possible to photograph from magnifications of 0.45 times to 16.5 times.

Circle No. 111 on Reader Service Card.



Close-up work with the photomakroskop



The Guinier camera from Philips

● A new Guinier camera is now being produced by Philips for use in X ray diffraction. The camera may be used either independently or in conjunction with a generator such as the Philips table-top PW 1729. It can also be used with a Philips vertical diffractometer, since it can be fitted to the opposite line focus of the system's X ray tube tower. The camera is provided in the form of a separate monochromator, camera housing and base units, which reduces any interactive effects during installation and operation. A lightweight housing holds both the specimen and film chamber. The specimen holder is factory-mounted at 45° to the incident beam axis, giving a full 100° (2θ) per 1 mm film length. The equipment uses narrow line widths to increase definition. Various accessories are available and the chamber may be operated under vacuum when handling easily oxidizable or hygroscopic samples. Plug-in specimen holders are also available.

Circle No. 112 on Reader Service Card

● A system is now available from Wild Heerbrugg for use in close-up work which is too small for the ordinary camera to be used alone, but is too large for the visual field of a microscope. This photomakroskop has a photographic range of 1:1 to 60:1. It has a macrozoom objective and a working distance of 42 to 188 mm.

Circle No. 113 on Reader Service Card.

● A new system for use in UV/visible photography has been launched by Fotodyne. The Fotosystem 1000 is a table-top system which comes with a floor-standing base. Its light sockets will accommodate bulbs of 254, 300 and 366 nm wavelengths and also white light bulbs. The system has an extra accessory which allows the standard Polaroid film system to be changed to accommodate other films such as 35 mm, instant 8×10 inch colour and those for close-up copy work.

Circle No. 114 on Reader Service Card

Accessories

● A thermoplastic recording film with improved spectral sensitivity at the ruby laser wavelength (694.3 nm) is now being offered by Newport. Holographic diffraction efficiencies of 10 % can be obtained with exposures of 10 erg per cm² for holograms.

Circle No. 115 on Reader Service Card.

● Wild Heerbrugg has developed a series of lenses for use in photogrammetry (the use of aerial photographs in surveying and mapmaking). The lenses can be used with the Avioptot RC10 and the RC10A cameras and cover the range 400–900nm.

Circle No. 116 on Reader Service Card

● Polaroid now offer two systems for taking close-up photographs. An adaptation of the Polaroid CU-5 camera with built-in lighting, fixed focus, pre-calculated exposure settings and object-framing capacities is designed for use in, for example, intra-oral dental records. The second system, the CU-70 close-system, allows an SX-70 camera to produce 1:1, 2:1 or 3:1 colour photographs. Three models of SX-70 cameras can be used with the system — the Autofocus, the Sonar and the Alpha (models 1 or 2).

Circle No. 117 on Reader Service Card

ADVERTISEMENTS

**Do you need
HEAVY WATER?**

For lowest prices on D₂O call (415) 234-4130.

BIO-RAD

Bio-Rad Laboratories
2200 Wright Ave.
Richmond, CA 94804



THERMAL STAGE

Model TS-2 holds any temperature, minus 25°C to plus 60°C. It's thermoelectric, has no moving parts, no vibration.

The TS-2 stage fits most microscopes. Temperature control is automatic and the equipment includes a microprobe thermometer.

BAILEY INSTRUMENTS INC.
Saddle Brook, N.J. 07662 • U.S.A.

FOR THE MOST VALUE FOR YOUR RESEARCH DOLLAR,
ASK FOR OUR NEW CATALOGUE

CUSTOM ANTISERUMS IN
RABBITS
and
GOATS

A =

A1 Quality

CALL
OR WRITE

TOM WICKS or MIKE SMITH

IN CA (916) 758-4300 • (800) 824-8540



antibodies incorporated

P.O. Box 1560 • Davis, CA 95617

● Edax has updated the Econ II windowless detecting unit for light element microanalysis of specimens whilst under examination with scanning electron microscopes. The new Econ III detecting unit both extends the analytical range to include boron and allows the detection of all elements with molecular weights less than or equal to uranium.

Circle No. 118 on Reader Service Card.

● Kontes now produces a multi-media fibre optic scanner which can be used to look at thin layer chromatography plates, paper, glass, slab or tube gels, textiles, metals, powders and films. The model 800 operates in both reflectance and transmission modes, and the scans are reproducible to $\pm 0.25\%$.

Circle No. 119 on Reader Service Card.

Microscopes

● The BH-2 system microscope series has recently been released by Olympus. A range of attachments is also available for the new microscopes, including a phase contrast attachment, and the BH2-PC, for observing unstained or living material.

Circle No. 120 on Reader Service Card.

● A new automatic image analyser has been developed by the ophthalmic instruments division of Reichert-Jung. It can be interfaced with most scanning electron microscopes (SEMs) or scanning transmission electron microscopes (STEMs). The new system, called SEM-IPS, can segment images from SEMs irrespective of their complexity or the background noise. The system is software-based and requires no additional hardware. It provides for up to four video signals in parallel to be stored in the digital framestore, and X ray counts from up to eight element windows can also be stored in parallel for energy dispersive analysis (EDX) or wavelength dispersive analysis (WDX). The digital scan generator has a 12 bit (4,096 × 4,096 pixel) resolution with a 4 μsecond per pixel dwell time. Framestore capacity is available to a maximum of 16 Mbytes.

Circle No. 121 on Reader Service Card.

● The Philips SEM 505 scanning electron microscope system is designed to facilitate laboratory automation. It has a new computer interface, the PW 6791/00, which provides interactive communication between the microscope and the built-in DEC LSI-11 computer of an Edax energy-dispersive X ray analyser, or a separate DEC PDP 11 computer. The SEM 505 incorporates a degree of intelligence with a 'data link' bus system for communication between various control modules. This new interface adds a decision-making capability to solve application problems. For applications such as chemical analysis, elemental mapping and image analysis, the full Edax library of software is available.

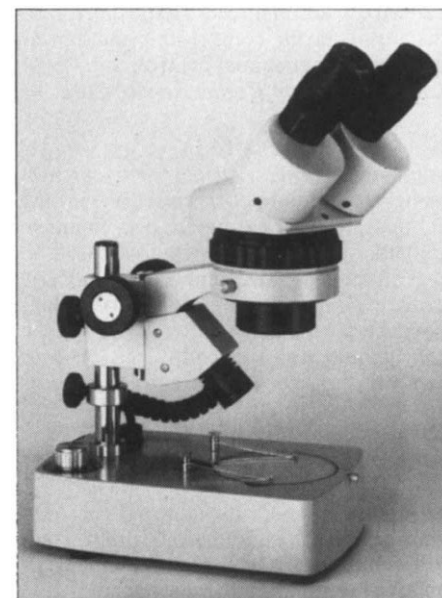
Circle No. 122 on Reader Service Card.

● The Kodavue electrophoresis visualization kit, which is for use with proteins and polypeptides separated on polyacrylamide gels, has been introduced by Eastman Kodak Company. The new kit allows proteins and polypeptides separated on polyacrylamide gels to be visualized in one hour, and provides a sensitivity of 1–2 ng. Sensitization of the protein band with the Kodavue kit occurs at the gel surface so that a single procedure applies to all gel thicknesses up to 1.5 mm and concentrations up to 20%. Only small volumes of reagents are required. The reagents featured in the kit are stable at room temperature and do not require refrigeration, and the Kodavue kit has a shelf life of several years. It contains sufficient reagents to visualize up to 24 14 × 16 cm gels.

Circle No. 120 on Reader Service Card

● Vickers Instruments have extended their range of stereoscopic microscopes for laboratory and general industrial applications. The new 3D-M range is comprised of wide-field stereo and zoom stereo microscopes. There are two 3D-M zoom instruments, the 1 × to 3 × or the 0.8 × to 4.5 ×, together with 10 ×, 15 × or 20 × eyepieces and supplementary 1.5 × or 2 × lenses, gives an overall magnification range of 8 × up to 180 ×. The working distance of the microscopes is 90 mm in the absence of supplementary lenses. The range of non-zoom, widefield stereo microscopes includes fixed-focus 1 × and 2 × models, plus dual power 1 × and 2 ×, or 1 × and 3 × options. A range of 10 ×, 15 × and 20 × eyepieces with supplementary lenses extend the selection of instruments from fixed magnification models to more versatile multi-magnification units. A wide range of accessories, including polarizing stages, dark ground facilities, gliding stages, specimen clamps and photographic attachments, are also available.

Circle No. 123 on Reader Service Card.



Vickers' new stereomicroscope